

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: August 31, 2003, 16:33:11 ; Search time 85 Seconds  
(without alignments)  
1092.412 Million cell updates/sec

Title: US-09-833-041-18

Perfect score: 3103  
Sequence: 1 DAHKEVAHREFKIDGENFK.....TCFAEKGKLVAAASQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq 19Jun03.\*  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	3103	100.0	585	10	AAP90388	Mature human serum
2	3103	100.0	585	11	AA05318	Human serum albumi
3	3103	100.0	585	11	AA08457	Human serum albumi
4	3103	100.0	585	16	AA08301	Human serum albumi
5	3103	100.0	585	18	AA02011	HSA protein sequen
6	3103	100.0	585	21	AA04873	Amino acid sequenc
7	3103	100.0	585	21	AA08394	yeast codon-biased
8	3103	100.0	585	22	AB073006	Human mature album
9	3103	100.0	585	22	AA013399	Human albumin (HA)

10	3103	100.0	585	22	AA052567	Mature human serum
11	3103	100.0	585	22	AA013129	Human albumin (HA)
12	3103	100.0	585	22	AA012403	Human albumin (HA)
13	3103	100.0	585	22	AA08578	Human serum albumi
14	3103	100.0	585	23	AB071291	Glycosylated prote
15	3103	100.0	585	23	AB063321	Human serum albumi
16	3103	100.0	585	23	AB000986	B lymphocyte stimu
17	3103	100.0	585	23	AB033847	Human B lymphocyte
18	3103	100.0	585	23	AA075220	Mature form of hum
19	3103	100.0	609	21	AA036542	Recombinant human
20	3103	100.0	609	21	AA036549	Recombinant human
21	3103	100.0	609	21	AA078147	Pre human serum al
22	3103	100.0	609	24	AB057252	Human serum albumi
23	3103	100.0	609	24	AB057253	Human serum albumi
24	3103	100.0	610	14	AA039510	Chimeric human ser
25	3103	100.0	616	24	AA030916	Val8-GLP-1-human s
26	3103	100.0	624	24	AA030919	Human serum albumi
27	3103	100.0	631	24	AA030917	Val8-GLP-1-linker-
28	3103	100.0	640	24	AA030918	Val8-GLP-2-GLP-1-C
29	3103	100.0	640	24	AA030920	Exendin-4-linker-h
30	3103	100.0	670	21	AA036543	Recombinant human
31	3103	100.0	670	21	AA036550	Recombinant human
32	3103	100.0	783	14	AA039473	Prepro-HSA-G-CSF c
33	3103	100.0	787	14	AA039477	G-CSF-(GLY)4-HSA c
34	3103	100.0	853	14	AA039472	HSA-vWF(470-713) f
35	3099	99.9	585	10	AA093344	Sequence of mature
36	3099	99.9	585	19	AA059841	Mature protein of
37	3099	99.9	608	17	AA096229	Human serum albumi
38	3099	99.9	609	14	AA030089	Sequence of human
39	3099	99.9	609	17	AA096232	Human serum albumi
40	3099	99.9	609	17	AA094572	Cancer metastasis
41	3099	99.9	609	17	AA088913	Human serum albumi
42	3099	99.9	609	19	AA048095	Human serum albumi
43	3099	99.9	609	20	AA006994	Human albumin. Ho
44	3099	99.9	609	22	AA004148	Myosin light chain
45	3099	99.9	609	23	AB032802	Human serum albumi

## ALIGNMENTS

### RESULT 1

AA090388  
ID AAP90388 standard; protein: 585 AA.  
XX  
AC AAP90388;  
XX  
DT 25-MAR-2003 (updated)  
DT 01-NOV-1989 (first entry)  
XX  
DE Mature human serum albumin polypeptide.  
XX  
KW Human serum albumin; mature protein; new polypeptides;  
KW plasma expanders.  
XX  
OS Homo sapiens (Human).  
XX  
PN EP322094-A.  
XX  
PD 28-JUN-1989.  
XX  
PF 25-OCT-1988; 88EP-0310000.  
XX  
PR 30-OCT-1987; 87GB-0025529.  
XX  
(DELEZ ) DELTA BIOTECHNOLOGY LTD.  
XX  
PI Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;  
XX  
DR WPI, 1989-186464/26.  
XX  
DR N-PSDB; AA090128.  
XX  
PT New N-terminal fragments of human serum albumin

PT - esp. useful as blood plasma expanders.

PS Disclosure; fig 2; 20pp; English.

XX Mature protein of human serum albumin (see corresp. AAN90128).

CC Used to make new N-terminal fragments which are used as plasma

CC expanders, or as substitutes for HSA or BSA, in tissue culture

CC media.

CC (updated on 25-MAR-2003 to correct PA field.)

XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 10; Length 585;

Best Local Similarity 100.0%; Pred. No. 9.5e-255;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIATAFYAQYLOQCPFFEDHVKLVNEVTEFAKTCVADSAE 60

DB 1 DAHSEVAHFRKDLGEENFKALVLIATAFYAQYLOQCPFFEDHVKLVNEVTEFAKTCVADSAE 60

QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPALPRIVPEV 120

DB 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPALPRIVPEV 120

QY 121 DVMCTAFHDNEETFLKKYLYETARRHPFYAPPELLFFAKRYKAATFECQQAADKAACLLP 180

DB 121 DVMCTAFHDNEETFLKKYLYETARRHPFYAPPELLFFAKRYKAATFECQQAADKAACLLP 180

QY 181 KLDELREGKASSAKORLKASLOKFGERAFKAWARLSQRPKAEFAEVSCLVTDLT 240

DB 181 KLDELREGKASSAKORLKASLOKFGERAFKAWARLSQRPKAEFAEVSCLVTDLT 240

QY 241 VHTCCGHDLECCADRADLAKYICENQDSISSKKECCCKPLEKSHCIAEVENDEMPA 300

DB 241 VHTCCGHDLECCADRADLAKYICENQDSISSKKECCCKPLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNTAEAKDVFLGMFLYETARRHPDYSVLLILRLAKTYETILEK 360

DB 301 DLPSLAADFVESKDVCKNTAEAKDVFLGMFLYETARRHPDYSVLLILRLAKTYETILEK 360

QY 361 CAADAPHECYAKVDFEFKPLVEEPONLIKONCELFQELGEYKFNALLVRYTKKVPQVST 420

DB 361 CAADAPHECYAKVDFEFKPLVEEPONLIKONCELFQELGEYKFNALLVRYTKKVPQVST 420

QY 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

DB 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

QY 481 LVNRRPCFSALEVDETYVPKEFNAETFFHADICTLSEKERQIKKQTALVELVHKPKAT 540

DB 481 LVNRRPCFSALEVDETYVPKEFNAETFFHADICTLSEKERQIKKQTALVELVHKPKAT 540

QY 541 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEBGKKLVAAASQAALGL 585

DB 541 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEBGKKLVAAASQAALGL 585

RESULT 2

ID AAR05318

XX AAR05318 standard; protein; 585 AA.

AC AAR05318;

XX 08-OCT-1990 (first entry)

XX Human serum albumin gene product.

XX Human serum albumin; HSA-A; yeast; ds.

XX Homo sapiens.

XX JP02117384-A.

XX

PD 01-MAY-1990.

XX 26-OCT-1988; 88JP-0268302.

XX 26-OCT-1988; 88JP-0268302.

XX (TOFU ) TOA NENRYO KOGYO KK.

XX WPI; 1990-176228/23.

XX N-PSDB; AAQ04719.

XX Human serum albumin prepn. by yeast host -

PT by culturing transformed plasmid yeast to produce serum, and

PT removing it.

XX Disclosure; Page ?; ?pp; Japanese.

CC Mature HSA-A may be produced using the sequence incorporated into a

CC plasmid vector with suitable controllers, and transferred to a yeast

CC expression system.

XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 11; Length 585;

Best Local Similarity 100.0%; Pred. No. 9.5e-255;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIATAFYAQYLOQCPFFEDHVKLVNEVTEFAKTCVADSAE 60

DB 1 DAHSEVAHFRKDLGEENFKALVLIATAFYAQYLOQCPFFEDHVKLVNEVTEFAKTCVADSAE 60

QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPALPRIVPEV 120

DB 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPALPRIVPEV 120

QY 121 DVMCTAFHDNEETFLKKYLYETARRHPFYAPPELLFFAKRYKAATFECQQAADKAACLLP 180

DB 121 DVMCTAFHDNEETFLKKYLYETARRHPFYAPPELLFFAKRYKAATFECQQAADKAACLLP 180

QY 181 KLDELREGKASSAKORLKASLOKFGERAFKAWARLSQRPKAEFAEVSCLVTDLT 240

DB 181 KLDELREGKASSAKORLKASLOKFGERAFKAWARLSQRPKAEFAEVSCLVTDLT 240

QY 241 VHTCCGHDLECCADRADLAKYICENQDSISSKKECCCKPLEKSHCIAEVENDEMPA 300

DB 241 VHTCCGHDLECCADRADLAKYICENQDSISSKKECCCKPLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNTAEAKDVFLGMFLYETARRHPDYSVLLILRLAKTYETILEK 360

DB 301 DLPSLAADFVESKDVCKNTAEAKDVFLGMFLYETARRHPDYSVLLILRLAKTYETILEK 360

QY 361 CAADAPHECYAKVDFEFKPLVEEPONLIKONCELFQELGEYKFNALLVRYTKKVPQVST 420

DB 361 CAADAPHECYAKVDFEFKPLVEEPONLIKONCELFQELGEYKFNALLVRYTKKVPQVST 420

QY 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

DB 421 PTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

QY 481 LVNRRPCFSALEVDETYVPKEFNAETFFHADICTLSEKERQIKKQTALVELVHKPKAT 540

DB 481 LVNRRPCFSALEVDETYVPKEFNAETFFHADICTLSEKERQIKKQTALVELVHKPKAT 540

QY 541 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEBGKKLVAAASQAALGL 585

DB 541 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEBGKKLVAAASQAALGL 585

RESULT 3

ID AAR08457

XX AAR08457 standard; Protein; 585 AA.

AC AAR08457;

```

XX
DT 25-MAR-2003 (updated)
DT 16-APR-1991 (first entry)
XX
XX Human serum albumin.
XX
XX HSA; folding; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 123..303
XX /label= A
XX
XX Region 1..303
XX /label= B
XX Region 123..585
XX /label= C
XX
XX JP02227079-A.
XX
XX 10-SEP-1990.
XX
XX 25-AUG-1989; 89JP-0217540.
XX
XX 25-AUG-1989; 89JP-0217540.
XX
XX (TOFU ) TONEN CORP.
XX
XX WPI; 1990-317325/42.
XX N-PSDB; AAQ06099.
XX
XX New human serum albumin fragments - used to bond medicines and for
XX stable folding of protein(s).
XX
XX Claim 1; Fig 8; 24pp; Japanese.
XX
XX Fragments A-C of HSA are expressed as fusion proteins with the
XX signal peptide of E. coli alkaline phosphatase. The fragments are
XX selected for their specific properties. The C-terminal truncated
XX fragment, B, does not bind long-chain fatty acids but does bind to
XX various medicines at the central region. The N-terminal truncated
XX fragment, C, has good stability in protein folding. The central
XX segment, A, has characteristics of both B and C.
XX See also AAQ06099-C06098.
XX (Updated on 25-MAR-2003 to correct PD field.)
XX (Updated on 25-MAR-2003 to correct PF field.)
XX (Updated on 25-MAR-2003 to correct PR field.)
XX
XX Sequence 585 AA;
XX
XX Query Match 100.0%; Score 3103; DB 11; Length 585;
XX Best Local Similarity 100.0%; Pred. No. 9.5e-255;
XX Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 DAHSEVAHFKDLGENFKALVLIAPVLAQYLCQCPEDHVKLVNEVTEFAKTCVADSEAE 60
DB 1 DAHSEVAHFKDLGENFKALVLIAPVLAQYLCQCPEDHVKLVNEVTEFAKTCVADSEAE 60
XX
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVIRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVIRPEV 120
XX
QY 121 DVMCTAFHDNEETFLKKYIETARRHPYFYAPPELLFPKRYKAATECCQAADKAACILP 180
DB 121 DVMCTAFHDNEETFLKKYIETARRHPYFYAPPELLFPKRYKAATECCQAADKAACILP 180
XX
QY 181 KIDLELRDEGKASSAKORLKASLOKFGERAFKAWAVARLSORPPKAEFYSKVLDTLTK 240
DB 181 KIDLELRDEGKASSAKORLKASLOKFGERAFKAWAVARLSORPPKAEFYSKVLDTLTK 240
XX
QY 241 VHTCECHGDLLEACDADRADIACYICENODSISSEKCECKPILLEKSHCIAEVENDMPA 300
DB 241 VHTCECHGDLLEACDADRADIACYICENODSISSEKCECKPILLEKSHCIAEVENDMPA 300

```

```

QY 301 DLPSLAADFVESKDVCKKNYAERAKDVFGLMFLEYEARHPDYSVLLRLAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKKNYAERAKDVFGLMFLEYEARHPDYSVLLRLAKTYETTTLEKC 360
XX
QY 361 CAADPHECYAKVDFEFKPLVEEPQNLKONCELPOLGEYKFNALLVRYTKVPQVST 420
DB 361 CAADPHECYAKVDFEFKPLVEEPQNLKONCELPOLGEYKFNALLVRYTKVPQVST 420
XX
QY 421 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEADYLSVVLNOLCVLHKEKTPVSDRYTKCCTES 480
DB 421 PTLVEVSRLNGKVGSKCKCKHPEAKRMPCAEADYLSVVLNOLCVLHKEKTPVSDRYTKCCTES 480
XX
QY 481 LVNRRPCFSALEVDYETVYKPFNAETFFHADICTLSEKERQIKQALVELVYKHPKAT 540
DB 481 LVNRRPCFSALEVDYETVYKPFNAETFFHADICTLSEKERQIKQALVELVYKHPKAT 540
XX
QY 541 KEQLKAVMDDFAAAFVEKCKCKADDDKTCFAEEGKKLVAAASQAALGL 585
DB 541 KEQLKAVMDDFAAAFVEKCKCKADDDKTCFAEEGKKLVAAASQAALGL 585

```

## RESULT 4

AAR80301  
ID AAR80301 standard; Protein; 585 AA.

XX AC AAR80301;

XX XX

DT 25-MAR-2003 (updated)

DT 17-JAN-1996 (first entry)

XX XX Human serum albumin.

XX XX Serum albumin; HSA; aspartyl protease-3; Yap3p;

KW Saccharomyces cerevisiae.

XX OS Homo sapiens.

XX PN W09523857-A1.

XX PD 08-SEP-1995.

XX PF 01-MAR-1995; 95WO-GB00434.

XX PR 05-MAR-1994; 94GB-0004270.

XX PA (DELZ ) DELTA BIOTECHNOLOGY LTD.

XX PI Kerrywilliams SM, Gilbert SC;

XX XX

DR WPI; 1995-320572/41.

DR N-PSDB; AAQ98695.

XX Yeast with reduced levels of aspartyl protease 3 proteolytic

PT activity - used to secrete human albumin without prodn. of the 45

PT KD fragment

XX Example 1; Page 26-28; 50pp; English.

XX CC The cDNA given in AA098695, which encodes HSA (AAR80301), was subjected

CC to site-directed mutagenesis to investigate the role of

CC endoproteases in the generation of a 45 kDa albumin fragment obtd.

CC when the cDNA is expressed in S. cerevisiae. Mutations were: R410A;

CC L407A, L408V, V409A; and R410A, K413Q, K414Q. The latter set of

CC mutations, especially, improved stability of HSA to yeast Yap3p

CC proteolytic cleavage, allowing increased prodn. of recombinant HSA.

CC (Updated on 25-MAR-2003 to correct PI field.)

XX XX Sequence 585 AA;

SQ

Query Match 100.0%; Score 3103; DB 16; Length 585;

Best Local Similarity 100.0%; Pred. No. 9.5e-255;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREKDLGEENFKALVLIATAQYLOQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60  
 DB 1 DAHSEVAHREKDLGEENFKALVLIATAQYLOQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60  
 QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVREVEV 120  
 DB 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVREVEV 120  
 QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFAKRYKAAPTECCQAADKAACLLP 180  
 DB 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFAKRYKAAPTECCQAADKAACLLP 180  
 QY 181 KLDELDRSGKASSAKQRLKASLOKFGERAFKANAVARLSORFPKAEFAEVSKLVTDLTK 240  
 DB 181 KLDELDRSGKASSAKQRLKASLOKFGERAFKANAVARLSORFPKAEFAEVSKLVTDLTK 240  
 QY 241 VITECHGDLLECADRADLAKYICENQDSISSKLKCECKPILLESKSHCIAEVENDEMPA 300  
 DB 241 VITECHGDLLECADRADLAKYICENQDSISSKLKCECKPILLESKSHCIAEVENDEMPA 300  
 QY 301 DLPSLAADFVSKDCKVKNYAEAKDVFGLMFLYEYARRHPDYSVVLRLRLAKTYETLEKC 360  
 DB 301 DLPSLAADFVSKDCKVKNYAEAKDVFGLMFLYEYARRHPDYSVVLRLRLAKTYETLEKC 360  
 QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKONCELFGEQYKFNQALLVRYTKKVPQYST 420  
 DB 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKONCELFGEQYKFNQALLVRYTKKVPQYST 420  
 QY 421 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
 DB 421 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
 QY 481 LVNRRPCFSALEVDVETVVPKEFNAETFFHADICTLSEKEROIKKOTALVELVKKHPRAT 540  
 DB 481 LVNRRPCFSALEVDVETVVPKEFNAETFFHADICTLSEKEROIKKOTALVELVKKHPRAT 540  
 QY 541 KBQLKAVMDDFAAAFVEKCKKADDKETCFPAEKGKLVAAASQAALGL 585  
 DB 541 KBQLKAVMDDFAAAFVEKCKKADDKETCFPAEKGKLVAAASQAALGL 585

## RESULT 5

AAO20111  
 ID AAO20111 standard; Protein; 585 AA.

AC AAO20111;

XX AAO20111;

DT 06-AUG-2002 (first entry)

XX HSA protein sequence related to the growth hormone protein.

DE Serum albumin-growth hormone fusion protein; growth hormone;

XX Down's syndrome.

XX Unidentified.

XX KR99076789-A.

XX 15-OCT-1999.

XX 25-JUN-1998; 98KR-0704914.

XX 30-DEC-1995; 95GB-0026733.

XX 19-DEC-1996; 96WO-GB03164.

XX (DELTA ) DELTA BIOTECHNOLOGY LTD.

XX WPI; 1997-363680/55.

XX N-PSDB; AAK99568.

XX Serum albumin-growth hormone fusion protein - useful to treat growth

PT hormone related diseases, e.g. Down's syndrome

XX Disclosure; Fig 6; 2lpp; Korean.  
 XX The invention relates to a serum albumin-growth hormone fusion protein -  
 CC useful to treat growth hormone related diseases such as Down's syndrome.  
 CC This sequence represents a HSA protein related to the serum albumin-  
 CC growth hormone protein of the invention.

SQ Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 18; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-255;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREKDLGEENFKALVLIATAQYLOQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60  
 DB 1 DAHSEVAHREKDLGEENFKALVLIATAQYLOQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60  
 QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVREVEV 120  
 DB 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVREVEV 120  
 QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFAKRYKAAPTECCQAADKAACLLP 180  
 DB 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFAKRYKAAPTECCQAADKAACLLP 180  
 QY 181 KLDELDRSGKASSAKQRLKASLOKFGERAFKANAVARLSORFPKAEFAEVSKLVTDLTK 240  
 DB 181 KLDELDRSGKASSAKQRLKASLOKFGERAFKANAVARLSORFPKAEFAEVSKLVTDLTK 240  
 QY 241 VITECHGDLLECADRADLAKYICENQDSISSKLKCECKPILLESKSHCIAEVENDEMPA 300  
 DB 241 VITECHGDLLECADRADLAKYICENQDSISSKLKCECKPILLESKSHCIAEVENDEMPA 300  
 QY 301 DLPSLAADFVSKDCKVKNYAEAKDVFGLMFLYEYARRHPDYSVVLRLRLAKTYETLEKC 360  
 DB 301 DLPSLAADFVSKDCKVKNYAEAKDVFGLMFLYEYARRHPDYSVVLRLRLAKTYETLEKC 360  
 QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKONCELFGEQYKFNQALLVRYTKKVPQYST 420  
 DB 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKONCELFGEQYKFNQALLVRYTKKVPQYST 420  
 QY 421 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
 DB 421 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
 QY 481 LVNRRPCFSALEVDVETVVPKEFNAETFFHADICTLSEKEROIKKOTALVELVKKHPRAT 540  
 DB 481 LVNRRPCFSALEVDVETVVPKEFNAETFFHADICTLSEKEROIKKOTALVELVKKHPRAT 540  
 QY 541 KBQLKAVMDDFAAAFVEKCKKADDKETCFPAEKGKLVAAASQAALGL 585  
 DB 541 KBQLKAVMDDFAAAFVEKCKKADDKETCFPAEKGKLVAAASQAALGL 585

## RESULT 6

AAO20111

ID AAO20111 standard; protein; 585 AA.

AC AAO20111;

XX AAO20111;

DT 08-AUG-2000 (first entry)

XX Amino acid sequence of a human albumin protein.

DE Human; albumin; ischemic state; serum protein; metal ion salt;

XX peroperative ischemia; ischemia; myocardial infarction;

XX progressive coronary artery disease.

XX Homo sapiens.

XX Key Location/Qualifiers

PH Modified-site 1

PT Serum albumin-growth hormone fusion protein - useful to treat growth  
 hormone related diseases, e.g. Down's syndrome



PT /note= "optionally acetylated, and claimed under  
FT claim 56"  
PN WO200020840-A1.  
XX 13-APR-2000.  
PD 01-OCT-1999; 99WO-US22905.  
XX 02-OCT-1998; 98US-0102738.  
PR 02-OCT-1998; 98US-0165581.  
PR 02-OCT-1998; 98US-0165926.  
PR 11-JAN-1999; 99US-0115392.  
XX (ISCH-) ISCHEMIA TECHNOLOGIES INC.  
PA Bar-Or D, Lau E, Winkler JV;  
XX WPI; 2000-303843/26.  
DR New method for the continuous detection of ischemic states comprises  
PT detecting and quantifying the existence of an alteration of the serum  
PT protein albumin -  
XX Disclosure; Page 97-100; 105pp; English.  
PS The present sequence represents human albumin protein. The specification  
XX describes a method for the continuous detection of ischemic states. The  
CC method comprises detecting and quantifying the existence of an alteration  
CC of the serum protein albumin. The method comprises contacting a  
CC biological sample containing albumin from the patient with an excess  
CC quantity of a metal ion salt, where the metal ion binds to the N-terminus  
CC of naturally occurring human albumin, to form a mixture containing bound  
CC metal ions and unbound metal ions, and then determining the amount of  
CC metal ions bound to the albumin N-terminus. The amount of bound metal  
CC ions is correlated to a known value to determine the occurrence or  
CC non-occurrence of an ischemic event. The methods are useful for detection  
CC of ischemic states. The methods are also useful for distinguishing  
CC pericardial ischemia from ischemia caused by , amongst other things,  
CC myocardial infarctions and progressive coronary artery disease.  
XX  
SQ Sequence 585 AA;  
Query Match 100.0%; Score 3103; DB 21; Length 585;  
Best Local Similarity 100.0%; Pred. No. 9.5e-255;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
QY 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQEPNERCEFLQHKDDNPNLPRLVREPV 120  
DB 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQEPNERCEFLQHKDDNPNLPRLVREPV 120  
QY 121 DVNCTAFHDNETFLKYLVEIARRHPYFAPELFFAKRYKAAFTCCQAAKACLLP 180  
DB 121 DVNCTAFHDNETFLKYLVEIARRHPYFAPELFFAKRYKAAFTCCQAAKACLLP 180  
QY 181 KIDELDEKASSAKQRLKASLOXTGERAFKAWAVARLSORFFKAEFAVSKLYTDTLK 240  
DB 181 KIDELDEKASSAKQRLKASLOXTGERAFKAWAVARLSORFFKAEFAVSKLYTDTLK 240  
QY 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLECECEKPLEKSHCIAEVENDEMPA 300  
DB 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLECECEKPLEKSHCIAEVENDEMPA 300  
QY 301 DLPSTLAADPVESKVCNKAFAKDFLGMFLYEAARRHPDYVSVLLLRILAKYETLEKC 360  
DB 301 DLPSTLAADPVESKVCNKAFAKDFLGMFLYEAARRHPDYVSVLLLRILAKYETLEKC 360  
QY 361 CAADPHECVAKYVDFEFPQNLKONCELFQDGEYKFNQALIVRYTKVPQVST 420

DB 361 CAADPHECVAKYVDFEFPQNLKONCELFQDGEYKFNQALIVRYTKVPQVST 420  
QY 421 PTLVEVSRNLGVGSKCKHPEAKRMPCAEYLSVLNOLCVLHEKTPVSDRVTKCTES 480  
DB 421 PTLVEVSRNLGVGSKCKHPEAKRMPCAEYLSVLNOLCVLHEKTPVSDRVTKCTES 480  
QY 481 LVNRRPCFSALVEVDYTVKBFNAETTFHADICTLSEKEROIKQKOTALVELVHKPKAT 540  
DB 481 LVNRRPCFSALVEVDYTVKBFNAETTFHADICTLSEKEROIKQKOTALVELVHKPKAT 540  
QY 541 KEQLKAVMDDFRAAFVEKCKKADKCTCFAEFGKKLVASQAALGL 585  
DB 541 KEQLKAVMDDFRAAFVEKCKKADKCTCFAEFGKKLVASQAALGL 585  
RESULT 7  
AAY83946  
ID AAY83946 standard; Protein; 585 AA.  
XX AAY83946;  
AC AAY83946;  
XX 28-JUL-2000 (first entry)  
DT Yeast codon-biased recombinant human serum albumin protein.  
XX Yeast codon-biased recombinant human serum albumin protein.  
XX Recombinant; human serum albumin; HSA; yeast codon bias; host cell;  
KW overlapping oligonucleotide; expression vector.  
XX Homo sapiens.  
OS Synthetic.  
XX CNL239103-A.  
XX 22-DEC-1999.  
XX 17-JUN-1998; 98CN-0102506.  
XX 17-JUN-1998; 98CN-0102506.  
PR (HALJ-) HALJI BIOENGINEERING CO LTD.  
XX Li S, Lu D;  
XX WPI; 2000-351198/31.  
DR N-PSDB; AAL0091.  
XX Process for preparing recombinant human serum albumin - which comprises  
PT yeast biased sex codons  
XX Disclosure; Fig 1; 44pp; Chinese.  
XX The method relates to a method of recombinantly producing human serum  
CC albumin (HSA) in yeast by altering the coding sequence of HSA to  
CC comprise a yeast codon bias. The complete HSA gene (AAL0091) was  
CC generated as three synthetic fragments (AAL0092-A10094) joined by  
CC recombinant DNA technology. Each HSA fragment was synthesized from  
CC overlapping oligonucleotide fragments that were extended. This sequence  
CC represents the complete sequence of the HSA encoded by the human gene  
CC with a yeast codon bias. The invention also covers a recombinant  
CC expression vector, yeast host cells carrying the recombinant expression  
CC vector and the process for producing human serum albumin in the yeast  
CC host cell, especially in secretory mode.  
XX Sequence 585 AA;  
Query Match 100.0%; Score 3103; DB 21; Length 585;  
Best Local Similarity 100.0%; Pred. No. 9.5e-255;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

```

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCAKOEPRNECFLOHKDDNPMLPLRVEY 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCAKOEPRNECFLOHKDDNPMLPLRVEY 120
QY 121 DVMCTAFHDNEETFLAKLYETARHPFFYAPELLFAKRYKAATFECQADKAACILP 180
Db 121 DVMCTAFHDNEETFLAKLYETARHPFFYAPELLFAKRYKAATFECQADKAACILP 180
QY 181 KLDELREGKASSAKORLKCSLQKGERAFKAWAVARLSORFPAEFAVSKLVTDLT 240
Db 181 KLDELREGKASSAKORLKCSLQKGERAFKAWAVARLSORFPAEFAVSKLVTDLT 240
QY 241 VHTCCGGDLLECADRADLAKYICENODSISSKLKECKECPLEKSHCIAEVENDEMPA 300
Db 241 VHTCCGGDLLECADRADLAKYICENODSISSKLKECKECPLEKSHCIAEVENDEMPA 300
QY 301 DPLSLAADFVSKDVKCYAEAKDVFGLGMFLYFARHPDYISVLLILAKTYETTTLEK 360
Db 301 DPLSLAADFVSKDVKCYAEAKDVFGLGMFLYFARHPDYISVLLILAKTYETTTLEK 360
QY 361 CAADAPHECYAKVDFEFKELVPEEPONLIKQNCLEFQGLGYKFNALLVRYTKKVPQVST 420
Db 361 CAADAPHECYAKVDFEFKELVPEEPONLIKQNCLEFQGLGYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLKGVSKCKKHPKAPCAEDYLSVNLQVLEHKTPTVSDRVTKKCTES 480
Db 421 PTLVEVSRNLKGVSKCKKHPKAPCAEDYLSVNLQVLEHKTPTVSDRVTKKCTES 480
QY 481 LVNRRPCFSALEVDYVYKFNATFTFHADICTLSEKROIKKOTALVELVKEKPKAT 540
Db 481 LVNRRPCFSALEVDYVYKFNATFTFHADICTLSEKROIKKOTALVELVKEKPKAT 540
QY 541 KEQLKAVMDFAAFVEKCKADKDKTCFAEKGKLVAAQAALGL 585
Db 541 KEQLKAVMDFAAFVEKCKADKDKTCFAEKGKLVAAQAALGL 585

RESULT 8
ABB79006
ID ID ABB79006 standard; Protein; 585 AA.
XX AC ABB79006;
XX DT 01-AUG-2002 (first entry)
XX DE Human mature albumin protein SEQ ID NO:18.
XX KW Human; growth hormone; hGH; albumin; human serum albumin; HSA;
KW albumin fusion protein; cytostatic; anorectic; immunosuppressive;
KW antidiabetic; antirheumatic; antiarthritic; psoriatic; cancer;
KW non-Hodgkin's lymphoma; obesity; type I diabetes mellitus; rheumatoid arthritis.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Domain 1..194
FT /label= 1
FT Domain 1..105
FT /label= subdomain
FT Disulfide-bond 53..62
FT Disulfide-bond 75..91
FT Disulfide-bond 90..101
FT Region 106..119
FT /note= "flexible inter-subdomain linker region"
FT Domain 120..194
FT /label= subdomain
FT Disulfide-bond 124..169
FT Disulfide-bond 168..177
FT Domain 195..387
FT /label= 2
FT Domain 195..291
FT /label= subdomain

```

```

FT Disulfide-bond 245..253
FT Disulfide-bond 265..279
FT Disulfide-bond 278..289
FT Region
FT /note= "flexible inter-subdomain linker region"
FT Domain 316..387
FT /label= subdomain
FT Disulfide-bond 316..361
FT Disulfide-bond 360..369
FT Domain 388..585
FT /label= 3
FT Domain 388..491
FT /label= subdomain
FT Disulfide-bond 392..438
FT Disulfide-bond 437..448
FT Disulfide-bond 461..477
FT Disulfide-bond 476..487
FT Region 492..511
FT /note= "flexible inter-subdomain linker region"
FT Domain 512..585
FT /label= subdomain
FT Disulfide-bond 514..559
FT Disulfide-bond 558..567
XX WC200179442-A2.
XX 25-OCT-2001.
XX 12-APR-2001; 2001WO-US11850.
XX 12-APR-2000; 2000US-229358P.
XX 25-APR-2000; 2000US-199384P.
XX 21-DEC-2000; 2000US-256931P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Haseltine WA;
XX WPI; 2001-611723/70.
XX N-PSDB; ABB87288.
XX New albumin fusion proteins, useful for treating diseases and disorders
XX such as cancer, comprise therapeutic protein fused to albumin -
XX Claim 1; Fig 11; 413pp; English.
XX The present invention describes an albumin fusion protein (I) comprising
XX a therapeutic protein: X and (a fragment or variant of) albumin
XX comprising a the fully defined sequence in ABB79006 of 585 amino acids,
XX (where the fragment or variant has albumin or therapeutic protein: X
XX activity). (I) can have cytostatic, anorectic, immunosuppressive,
XX antidiabetic, antirheumatic, antiarthritic and psoriatic activities.
XX Albumin fusion proteins are stabilised therapeutic proteins e.g.
XX antibodies to CS, C242 and C890 useful for treating various diseases
XX and disorders such as non-Hodgkin's lymphoma, cancer, obesity,
XX transplant rejection, type I diabetes mellitus, rheumatoid arthritis
XX and psoriasis. Fusing albumin to therapeutic proteins stabilises the
XX therapeutic protein, extends the shelf life and retains the in vitro or
XX in vivo biological activity. It also reduces the need to formulate
XX protein solutions with large excesses of carrier proteins to prevent
XX loss of therapeutic proteins due to factors such as binding to the
XX container. The fusion proteins are easily dispensed with a simple
XX formulation requiring minimal post storage manipulation. The fusion of
XX therapeutic proteins to albumin confers stability in aqueous or other
XX solution. The present sequence represents the mature human albumin (HA)
XX protein which is used in the exemplification of the present invention.
XX Sequence 585 AA;
XX Query Match 100.0%; Score 3103; DB 22; Length 585;
XX Best Local Similarity 100.0%; Pred. No. 9.5e-255;
XX Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 DAHSEVAHFKDLGEENFKALVLIAPAFYLOQCPEDHVKLVNEVTEFAKTCVADESSE 60  
 Db 1 DAHSEVAHFKDLGEENFKALVLIAPAFYLOQCPEDHVKLVNEVTEFAKTCVADESSE 60  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOBERNECFLOHKDNDNPLRLYRPEV 120  
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOBERNECFLOHKDNDNPLRLYRPEV 120  
 QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180  
 Db 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180  
 QY 181 KLDELDEGKASAKORLKCASLOKFGERAFKAWARLSQRFPAEFAVSKLYVLDLTK 240  
 Db 181 KLDELDEGKASAKORLKCASLOKFGERAFKAWARLSQRFPAEFAVSKLYVLDLTK 240  
 QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLAKECCERPLEKSHCHIAEVENDEMPA 300  
 Db 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLAKECCERPLEKSHCHIAEVENDEMPA 300  
 QY 301 DLPSTLAADFVESKDVCKNYAPAKDVFTLGMFLYIYARRHPDYSVLLLLRLAKTETYLEKC 360  
 Db 301 DLPSTLAADFVESKDVCKNYAPAKDVFTLGMFLYIYARRHPDYSVLLLLRLAKTETYLEKC 360  
 QY 361 CAADAPHECYAKVDFEPLVEEPQNLKONCELPOLGEYKFNQALIVRYTKKVPQVST 420  
 Db 361 CAADAPHECYAKVDFEPLVEEPQNLKONCELPOLGEYKFNQALIVRYTKKVPQVST 420  
 QY 421 PTLVEVSRNLGVSKCKCKHPEAKMPCAEYLSVNLQCVLHEKTPVSDRVTKCTES 480  
 Db 421 PTLVEVSRNLGVSKCKCKHPEAKMPCAEYLSVNLQCVLHEKTPVSDRVTKCTES 480  
 QY 481 LVNRRPCFSALEVDITYPKFENAEFTTFHADICTLSEKEROIKQFALVELVKHKPKAT 540  
 Db 481 LVNRRPCFSALEVDITYPKFENAEFTTFHADICTLSEKEROIKQFALVELVKHKPKAT 540  
 QY 541 KEQLKAWMDFAAFVEKCKADKDETCTFAEKGKLVAAASQAALGL 585  
 Db 541 KEQLKAWMDFAAFVEKCKADKDETCTFAEKGKLVAAASQAALGL 585

## RESULT 9

AAE13399

ID AAE13399 standard; Protein; 585 AA.

AC AAE13399;

XX 12-FEB-2002 (first entry)

DE Human albumin (HA) protein.

XX Human; albumin; HA; fusion protein; immune system disorder; syphilis;  
 KW transplant rejection; blood related disorder; myocardial infarction;  
 KW hyperproliferative disorder; acute myeloid leukemia; renal disorder;  
 KW glomerulonephritis; cardiovascular disorder; arrhythmia; rhinitis;  
 KW respiratory disorder; neurological disease; Alzheimer's disease;  
 KW endocrine disorder; pheochromocytoma; reproductive system disorder;  
 KW measles; gastrointestinal disorder; irritable bowel syndrome; HIV;  
 KW human immunodeficiency virus; wound healing; renal cell carcinoma;  
 KW melanoma; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Domain 54..61

XX Domain /label= Loop\_I

XX Domain 76..89

XX Domain /label= Loop\_II

XX Domain 92..100

XX Domain /label= Loop\_III

XX Domain 170..176

XX Domain /label= Loop\_IV

XX Domain 247..252

FT Domain /label= Loop\_V  
 266..277  
 FT /label= Loop\_VI  
 280..288  
 FT /label= Loop\_VII  
 362..368  
 FT /label= Loop\_VIII  
 439..447  
 FT /label= Loop\_IX  
 461..475  
 FT /label= Loop\_X  
 478..486  
 FT /label= Loop\_XI  
 560..566  
 FT /label= Loop\_XII

XX WO200179258-A1.

XX PN

XX PD 25-OCT-2001.

XX PF 12-APR-2001; 2001WO-US12008.

XX PR 12-APR-2000; 2000US-229358P.

XX PR 25-APR-2000; 2000US-199384P.

XX PR 21-DEC-2000; 2000US-256931P.

XX XX (HUMA-) HUMAN GENOME SCI INC.

XX PA (PRIN-) PRINCIPIA PHARM CORP.

XX XX Rosen CA, Sadeghi H, Prior CP, Turner AJ;

XX XX WPI; 2001-602931/68.

XX XX N-PSDB; RAD2287.

XX XX Albumin fusion proteins comprising a therapeutic protein and albumin,

XX XX useful in the treating metastatic renal cell carcinoma, metastatic

XX XX melanoma, malignant melanoma, renal cell carcinoma, HIV (human

XX XX immunodeficiency virus) or infection -

XX XX Claim 1; Fig 9; 325pp; English.

XX XX The invention relates to albumin fusion proteins comprising therapeutic

XX XX protein and human albumin (HA). The albumin fusion proteins are useful

XX XX in the treatment, prevention, diagnosis, and/or detection of diseases,

XX XX disorders such as immune system disorders (transplant rejection); blood

XX XX related disorders (myocardial infarction); hyperproliferative disorders

XX XX (childhood acute myeloid leukaemia); renal disorder (glomerulonephritis);

XX XX cardiovascular disorders (arrhythmias); respiratory disorders

XX XX (non-allergic rhinitis); neurological diseases (Alzheimer's disease);

XX XX endocrine disorders (pheochromocytoma); reproductive system disorders

XX XX (syphilis); infectious diseases (measles); gastrointestinal disorders

XX XX (irritable bowel syndrome) and wound healing. The albumin fusion

XX XX proteins are also used in the treatment of metastatic renal cell

XX XX carcinoma, metastatic melanoma, malignant melanoma and HIV (human

XX XX immunodeficiency virus) infection. Nucleic acid encoding albumin fusion

XX XX protein is useful in gene therapy. The present sequence is human

XX XX albumin (HA) protein.

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

Query Match 100.0%; Score 3103; DB 22; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-255;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDLGEENFKALVLIAPAFYLOQCPEDHVKLVNEVTEFAKTCVADESSE 60

Db 1 DAHSEVAHFKDLGEENFKALVLIAPAFYLOQCPEDHVKLVNEVTEFAKTCVADESSE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOBERNECFLOHKDNDNPLRLYRPEV 120

Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOBERNECFLOHKDNDNPLRLYRPEV 120

QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180

```

Db 121 DVMCTAFDNEETFLKKYLIEIARRHPYFAPELLFFAKRYKAAATECCQADRAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
QY 241 VHTCCGDLLECADRADLAKYICENODSISSSKLKCECKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCGDLLECADRADLAKYICENODSISSSKLKCECKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIYARRHPDYSVLLLRLLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIYARRHPDYSVLLLRLLAKTYETTLK 360
QY 361 CAADPHECYAKVDFEFKPLVEEPONLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
Db 361 CAADPHECYAKVDFEFKPLVEEPONLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRLGKVGSKCKKHPRKMPCAEDYLSVLNOLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRLGKVGSKCKKHPRKMPCAEDYLSVLNOLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYTVPKFNAETFFHADICTLSEKERQIKKOTALVELVHKPKAT 540
Db 481 LVNRRPCFSALEVDYTVPKFNAETFFHADICTLSEKERQIKKOTALVELVHKPKAT 540
QY 541 KEQLKAVMDDFAAFAVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAFAVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 10
AA052567
ID AA052567 standard; Protein; 585 AA.
AC AA052567;
XX
DT 05-FEB-2002 (first entry)
DE Mature human serum albumin.
XX
KW Human; serum albumin; HA; antiinflammatory; immunosuppressive; cardiant;
KW nootropic; neuroprotective; gene therapy; immune disorder; wound healing;
KW hyperproliferative disorder; renal disorder; cardiovascular disorder;
KW respiratory disorder; neurological disease; endocrine disorder;
KW reproductive system disorder; infectious disease;
KW gastrointestinal disorder.
XX
OS Homo sapiens.
XX
PN WO200179444-A2.
XX
PD 25-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US12013.
XX
PR 12-APR-2000; 2000US-229358P.
PR 25-APR-2000; 2000US-199384P.
PR 21-DEC-2000; 2000US-256931P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
Rosen CA, Haseltine WA;
XX
WPI: 2001-616755/71.
DR N-PSDB; ABA03057.
XX
Albunin fusion proteins comprising a therapeutic protein and albumin,
PT useful in the treating immune system disorders (e.g. transplant
PT rejection), blood related disorders (e.g. myocardial infarction) and
PT hyperproliferative disorders
XX

```

```

PS Claim 1: Fig 15; 606pp; English.
XX
CC The present invention relates to albumin fusion proteins, which comprise
CC a therapeutic protein and albumin. The present sequence is the protein
CC sequence for mature human serum albumin (HA), which was used to generate
CC the fusion proteins of the present invention. The albumin fusion proteins
CC are useful in the treatment, prevention, diagnosis, and/or detection of
CC diseases/disorders such as immune system disorders (e.g. myocardial
CC rejection), blood related disorders (e.g. childhood acute myeloid leukemia),
CC hyperproliferative disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g.
CC arrhythmias), respiratory disorders (e.g. non-allergic rhinitis),
CC neurological diseases (e.g. Alzheimer's disease), endocrine disorders
CC (e.g. pheochromocytoma), reproductive system disorders (e.g. syphilis),
CC infectious diseases (e.g. measles), gastrointestinal disorders (e.g.
CC irritable bowel syndrome) and wound healing.
XX
SQ Sequence 585 AA;
Query Match 100.0%; Score 3103; DB 22; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHREFKOLGGENFKALVIAFAQYLOQCPFFEDHVKLVNVEYTEFAKTCVADESAAE 60
Db 1 DAHSEVAHREFKOLGGENFKALVIAFAQYLOQCPFFEDHVKLVNVEYTEFAKTCVADESAAE 60
QY 61 NDCSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCEFLQHKDDNPPLRVLRPEV 120
Db 61 NDCSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNCEFLQHKDDNPPLRVLRPEV 120
QY 121 DVMCTAFDNEETFLKKYLIEIARRHPYFAPELLFFAKRYKAAATECCQADRAACLLP 180
Db 121 DVMCTAFDNEETFLKKYLIEIARRHPYFAPELLFFAKRYKAAATECCQADRAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
QY 241 VHTCCGDLLECADRADLAKYICENODSISSSKLKCECKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCGDLLECADRADLAKYICENODSISSSKLKCECKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIYARRHPDYSVLLLRLLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIYARRHPDYSVLLLRLLAKTYETTLK 360
QY 361 CAADPHECYAKVDFEFKPLVEEPONLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
Db 361 CAADPHECYAKVDFEFKPLVEEPONLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRLGKVGSKCKKHPRKMPCAEDYLSVLNOLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRLGKVGSKCKKHPRKMPCAEDYLSVLNOLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYTVPKFNAETFFHADICTLSEKERQIKKOTALVELVHKPKAT 540
Db 481 LVNRRPCFSALEVDYTVPKFNAETFFHADICTLSEKERQIKKOTALVELVHKPKAT 540
QY 541 KEQLKAVMDDFAAFAVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAFAVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 11
AAE13129
ID AAE13129 standard; Protein; 585 AA.
XX
AC AAE13129;
XX
DT 28-JAN-2002 (first entry)
XX
DE Human albumin (HA).

```

xx Human; albumin; HA; fusion protein; therapeutic protein; vulnary;  
 KW immune system disorder; transplant rejection; blood related disorder;  
 KW myocardial infarction; hyperproliferative disorder; glomerulonephritis;  
 KW childhood acute myeloid leukemia; cardiovascular disorder; arrhythmia;  
 KW respiratory disorder; gene therapy; non-allergic rhinitis; nontropic;  
 KW neurological disease; Alzheimer's disease; reproductive system disorder;  
 KW endocrine disorder; pheochromocytoma; infectious disease; antiarthritic;  
 KW measles; gastrointestinal disorder; irritable bowel syndrome; syphilis;  
 KW wound healing; antiinflammatory; immunosuppressive; neuroprotective;  
 KW cardiant; cytostatic; antileukemic; antirheumatic; antimicrobial;  
 KW renal disorder.  
 XX Homo sapiens.

xx Key Location/Qualifiers  
 FH Domain 54..61  
 FT /label= Loop\_I  
 FT Domain 76..89  
 FT /label= Loop\_II  
 FT Domain 92..100  
 FT /label= Loop\_III  
 FT Domain 170..176  
 FT /label= Loop\_IV  
 FT Domain 247..252  
 FT /label= Loop\_V  
 FT Domain 266..277  
 FT /label= Loop\_VI  
 FT Domain 280..288  
 FT /label= Loop\_VII  
 FT Domain 362..368  
 FT /label= Loop\_VIII  
 FT Domain 439..447  
 FT /label= Loop\_IX  
 FT Domain 461..475  
 FT /label= Loop\_X  
 FT Domain 478..486  
 FT /label= Loop\_XI  
 FT Domain 560..566  
 FT /label= Loop\_XII

xx WO200179443-A2.  
 PN  
 PD 25-OCT-2001.  
 XX 12-APR-2001; 2001WO-USL1924.  
 XX 12-APR-2000; 2000US-229358P.  
 PR 25-APR-2000; 2000US-199384P.  
 PR 21-DEC-2000; 2000US-256931P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA Rosen CA, Haseltine WA;  
 FI WPI; 2001-616754/71.  
 XX N-PSDB; AAD21638.  
 XX Albumin fusion proteins comprising a therapeutic protein and albumin,  
 PT useful in the treating immune system disorders (e.g. transplant  
 PT rejection), blood related disorders (e.g. myocardial infarction) and  
 PT hyperproliferative disorders -  
 XX  
 PT Claim 1; Fig 9; 38Opp; English.  
 PS  
 XX The invention relates to albumin fusion proteins comprising therapeutic  
 CC protein and human albumin (HA). Therapeutic protein fused to albumin  
 CC have an extended shelf-life. The albumin fusion proteins are useful in  
 CC the treatment, prevention, diagnosis and/or detection of diseases,  
 CC disorders such as immune system disorders (e.g. transplant rejection),  
 CC blood related disorders (e.g. myocardial infarction), hyperproliferative  
 CC disorders (e.g. childhood acute myeloid leukaemia), renal disorders  
 CC (e.g. glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),

CC respiratory disorders (e.g. non-allergic rhinitis), neurological  
 CC diseases (e.g. Alzheimer's disease), endocrine disorders (e.g.  
 CC pheochromocytoma), reproductive system disorders (e.g. syphilis),  
 CC infectious diseases (e.g. measles), gastrointestinal disorders (e.g.  
 CC irritable bowel syndrome) and wound healing. Nucleic acids encoding  
 CC albumin fusion protein is used in gene therapy. The present sequence  
 XX is human albumin (HA) protein.  
 SQ Sequence 585 AA;  
 Query Match 100.0%; Score 3103; DB 22; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 9,5e-255;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDLGEENFKALVLIAPAYLOCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
 DB 1 DAHSEVAHFKDLGEENFKALVLIAPAYLOCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHDDNPNLPRVPRV 120  
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHDDNPNLPRVPRV 120  
 QY 121 DVMTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLIP 180  
 DB 121 DVMTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLIP 180  
 QY 181 KLDELNDGKASSAKORLKCASQKGFGERAFKAWAVARLSORFFKAEFAVSKLVTDLTK 240  
 DB 181 KLDELNDGKASSAKORLKCASQKGFGERAFKAWAVARLSORFFKAEFAVSKLVTDLTK 240  
 QY 241 VHTCCGDLLESCADDDRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEYNDENMPA 300  
 DB 241 VHTCCGDLLESCADDDRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEYNDENMPA 300  
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVLLLRKATYETTTLEKC 360  
 DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVLLLRKATYETTTLEKC 360  
 QY 361 CAADPHECYAKVDFEFKPLVEEPQNLKONCELFQEGEYKFNALLVTKVPQVST 420  
 DB 361 CAADPHECYAKVDFEFKPLVEEPQNLKONCELFQEGEYKFNALLVTKVPQVST 420  
 QY 421 PTLVSVRNGLKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCCTES 480  
 DB 421 PTLVSVRNGLKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCCTES 480  
 QY 481 LVNRRPCFSALEVDYETYPKEFNAETTFHADICTLSEKEKQIKQKOTALVELVKHKKPAT 540  
 DB 481 LVNRRPCFSALEVDYETYPKEFNAETTFHADICTLSEKEKQIKQKOTALVELVKHKKPAT 540  
 QY 541 KEQLKAVMDFAAFVEKCKKADDDKTCFAEEGKKLVAAASQAALGL 585  
 DB 541 KEQLKAVMDFAAFVEKCKKADDDKTCFAEEGKKLVAAASQAALGL 585

RESULT 12  
 AAEL2403  
 ID AAEL2403 standard; Protein; 585 AA.  
 XX  
 AC AAEL2403;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Human albumin (HA).  
 XX Human; albumin; HA; immune system disorder; transplant rejection;  
 KW blood related disorder; myocardial infarction; glomerulonephritis;  
 KW hyperproliferative disorder; childhood acute myeloid leukaemia;  
 KW renal cell carcinoma; cardiovascular disorder; vulnary; melanoma;  
 KW arrhythmia; respiratory disorder; non-allergic rhinitis; antileukemic;  
 KW neurological disease; Alzheimer's disease; endocrine disorder; measles;  
 KW pheochromocytoma; reproductive system disorder; neuroprotective; syphilis;  
 KW infectious disease; gastrointestinal disorder; wound healing; nontropic;

irritable bowel syndrome; HIV; human immunodeficiency virus infection; cytostatic; antiinflammatory; gene therapy; immunosuppressive; cardiac; antiarthritic; antirheumatic; renal disorder; antimicrobial.

Homo sapiens.

Key Location/Qualifiers  
Domain 54..61  
/label= Loop\_I  
Domain 76..89  
/label= Loop\_II  
Domain 92..100  
/label= Loop\_III  
Domain 170..176  
/label= Loop\_IV  
Domain 247..252  
/label= Loop\_V  
Domain 266..277  
/label= Loop\_VI  
Domain 280..288  
/label= Loop\_VII  
Domain 362..368  
/label= Loop\_VIII  
Domain 439..447  
/label= Loop\_IX  
Domain 461..475  
/label= Loop\_X  
Domain 478..486  
/label= Loop\_XI  
Domain 560..566  
/label= Loop\_XII

WO2001179480-A1.

25-OCT-2001.

12-APR-2001; 2001WO-US11991.

12-APR-2000; 2000US-229358P.

25-APR-2000; 2000US-199384P.

21-DEC-2000; 2000US-256931P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Haseltine WA;

WPI: 2001-616756/71.

N-PSDB; ARD20005.

Albumin fusion proteins comprising a therapeutic protein and albumin, useful in the treating metastatic renal cell carcinoma, metastatic melanoma, malignant melanoma, renal cell carcinoma, HIV (human immunodeficiency virus) or infection -

Claim 1; Fig 9; 394pp; English.

The invention relates to human albumin (HA) fusion proteins and their corresponding nucleic acid sequences. Therapeutic proteins fused to albumin or its fragments have an extended shelf-life. The albumin fusion proteins are useful in the treatment, prevention, diagnosis, and/or detection of diseases, disorders such as immune system disorders (e.g. transplant rejection), blood related disorders (e.g. myocardial infarction), hyperproliferative disorders (e.g. childhood acute myeloid leukemia, metastatic renal cell carcinoma, metastatic melanoma, malignant melanoma, renal cell carcinoma), renal disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g. arrhythmias), respiratory disorders (e.g. non-allergic rhinitis), neurological diseases (e.g. Alzheimer's disease), endocrine disorders (e.g. pheochromocytoma), reproductive system disorders (e.g. syphilis), infectious diseases (e.g. measles), gastrointestinal disorders (e.g. irritable bowel syndrome), HIV (human immunodeficiency virus) infection and wound healing. Nucleic acids encoding albumin fusion protein is used in gene therapy. The present sequence is human albumin.

XX Sequence 585 AA;  
SQ Query Match 100.0%; Score 3103; DB 22; Length 585;  
Best Local Similarity 100.0%; Pred. No. 9.5e-255;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOCPPEDHVKLVNVEFTFAKTCVADESA 60  
DB 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOCPPEDHVKLVNVEFTFAKTCVADESA 60  
QY 61 NDKSLHTLFGDXLCTVATIRETYGEMADCCAKQEPERNECFLOHKDDNPMLPLVRPEV 120  
DB 61 NDKSLHTLFGDXLCTVATIRETYGEMADCCAKQEPERNECFLOHKDDNPMLPLVRPEV 120  
QY 121 DVMCTAFHDNEETFLKKYLYETARRHPYFYAPPELLFFAKYKAAFTCCQADKAACILP 180  
DB 121 DVMCTAFHDNEETFLKKYLYETARRHPYFYAPPELLFFAKYKAAFTCCQADKAACILP 180  
QY 181 KLDELDEGKASSAKORLKCASLQKFGRAFKAWARELSQRPKAFKAEVSKLVTDLTK 240  
DB 181 KLDELDEGKASSAKORLKCASLQKFGRAFKAWARELSQRPKAFKAEVSKLVTDLTK 240  
QY 241 VHTCCGHDLLCADDRLAKYICENQDSISSKLECCCKPLEKSHCIAFVENDEMPA 300  
DB 241 VHTCCGHDLLCADDRLAKYICENQDSISSKLECCCKPLEKSHCIAFVENDEMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360  
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFQNALIVRTTKVPQVST 420  
DB 361 CAAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFQNALIVRTTKVPQVST 420  
QY 421 PTLVEVSRNLGKVGSKCKKHPKAPCAEDYLSVNLQCLVHLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGKVGSKCKKHPKAPCAEDYLSVNLQCLVHLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDVETVPKFNATFTFHADICTLSEKERQIKKOTALVELVHKPKAT 540  
DB 481 LVNRRPCFSALEVDVETVPKFNATFTFHADICTLSEKERQIKKOTALVELVHKPKAT 540  
QY 541 KEQLKAVMDFAAFVEKCKCKADDDKTCFAERGKKLVAASQAALGL 585  
DB 541 KEQLKAVMDFAAFVEKCKCKADDDKTCFAERGKKLVAASQAALGL 585

RESULT 13

AAE08578

ID AAE08578 standard; Protein; 585 AA.

AC AAE08578;

XX 19-NOV-2001 (first entry)

DE Human serum albumin (HSA).

KW Human; albumin; cancer; cell proliferation; drug screening; biopsy.

XX Homo sapiens.

XX US6274305-B1.

XX 14-AUG-2001.

XX 19-DEC-1996; 96US-0769746.

XX 19-DEC-1996; 96US-0769746.

XX (TUFT ) UNIV TUFTS.

XX Sonnenschein C, Soto AM;

PI

XX WPI: 2001-540371/60.  
DR N-PSDB: AAD11488.  
XX  
XX  
XX Measuring human cell proliferation, useful in drug screening to  
PT determine the potential for inhibiting cancer cell proliferation and  
PT for evaluating biopsied tumors, comprises employing albumin-derived  
PT peptide -  
XX  
XX Claim 1: Fig 1: 20pp; English.  
XX  
XX The invention related to a method for testing cancer cells. The method is  
CC useful for measuring human cancer cell proliferation, particularly for  
CC determining the potential for inhibiting cancer cells proliferation using  
CC albumin-derived peptides. The invention is also useful for drug screening  
CC assays, as well as for evaluating biopsied tumors. The present sequence  
CC is human serum albumin (HSA) related to the invention.  
XX  
XX Sequence 585 AA;  
XX

Query Match	100.0%; Score 3103; DB 22; Length 585;
Best Local Similarity	100.0%; Pred. No. 9.5e-255;
Matches 585; Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	1 DAHKSVAHREKDLGEENFKALVLIAPAOYLQOCPEBHVKLIVNEVTEFAKTCVADSENE 60
DB	1 DAHKSVAHREKDLGEENFKALVLIAPAOYLQOCPEBHVKLIVNEVTEFAKTCVADSENE 60
QY	61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLPRVPRV 120
DB	61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLPRVPRV 120
QY	121 DVWCTAFHDNETFLFKKLYLFIARRHPHYFAPPELLFFAKRYKAATFCCOQADKAACLLP 180
DB	121 DVWCTAFHDNETFLFKKLYLFIARRHPHYFAPPELLFFAKRYKAATFCCOQADKAACLLP 180
QY	181 KLDELDEGKASSAQRLKCSAQKQGERAFKAWAVARLSQFFKAFBFAVSKLVDLTK 240
DB	181 KLDELDEGKASSAQRLKCSAQKQGERAFKAWAVARLSQFFKAFBFAVSKLVDLTK 240
QY	241 VHTECCHGDLLECADRRADLAKYTCENQDSISSKLKECCPKLEKSHCIAFVENDEMPA 300
DB	241 VHTECCHGDLLECADRRADLAKYTCENQDSISSKLKECCPKLEKSHCIAFVENDEMPA 300
QY	301 DLPSLAADFVESKDVCKNVAEAKDVFLGFMFLYEYARRHPDYSVLLRLAKTVEITLEKC 360
DB	301 DLPSLAADFVESKDVCKNVAEAKDVFLGFMFLYEYARRHPDYSVLLRLAKTVEITLEKC 360
QY	361 CAAADPHCEYAKVFDEFKPLVEEPQNLIKONCELFELQGEYKFNQALLVRYTKKVPQYST 420
DB	361 CAAADPHCEYAKVFDEFKPLVEEPQNLIKONCELFELQGEYKFNQALLVRYTKKVPQYST 420
QY	421 PTLVEYSRLGKVGSKCKCKEPAKRMPCBAEDYLSVVLNQLCVLHEKTPVSDRYTKCCIES 480
DB	421 PTLVEYSRLGKVGSKCKCKEPAKRMPCBAEDYLSVVLNQLCVLHEKTPVSDRYTKCCIES 480
QY	481 LVNRRPCFSALEVDEYTPYKPEFNAETFTFHADICTILSEKQRIKQOTALVELVKKPKAT 540
DB	481 LVNRRPCFSALEVDEYTPYKPEFNAETFTFHADICTILSEKQRIKQOTALVELVKKPKAT 540
QY	541 KEQLKAVMDDFRAFVEKCKCKADDKETCFABEGKKLVAASQAALGI 585
DB	541 KEQLKAVMDDFRAFVEKCKCKADDKETCFABEGKKLVAASQAALGI 585

RESULT 14	
ABG71291	
ID	ABG71291 standard; Protein; 585 AA.
XX	
AC	ABG71291;
XX	
DT	08-JAN-2003 (first entry)
XX	

DE	Glycosylated protein determination associated protein.
XX	
KW	Standard substance; accuracy control substance; glycosylated protein;
KW	glycosylated albumin; fructosamine; diabetes; antidiabetic.
XX	
OS	Unidentified.
XX	
FN	JF2002243731-A.
XX	
PD	28-AUG-2002.
XX	
PF	21-FEB-2001; 2001JP-0045085.
XX	
PR	21-FEB-2001; 2001JP-0045085.
XX	
PA	(KOKU-) KOKUSAI SHIYAKU KK.
PA	(YOSH ) YOSHITOMI PHARM IND KK.
XX	
DR	WPI; 2002-744850/81.
XX	
PT	A standard substance for determination of glycosylated protein
PT	including glycosylated albumin and fructosamine, used in diagnosis of
PT	diabetes
XX	
PS	Disclosure; Page 4; 6pp; Japanese.
XX	
CC	The present invention relates to a new standard and accuracy control
CC	substance for determination of glycosylated protein. The invention is
CC	useful for determination of glycosylated protein in the diagnosis of
CC	diabetes. Glycosylated albumin and fructosamine provide favourable
CC	dilution linearity. The present amino acid sequence represents the
CC	glycosylated protein determination associated protein as described in
CC	the invention.
XX	
SQ	Sequence 585 AA;
Query Match	100.0%; Score 3103; DB 23; Length 585;
Best Local Similarity	100.0%; Pred. No. 9.5e-255;
Matches 585; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 DAHSEVAHFRLDGBENFKALVLIAPAOYLQOCFFEDHVKLIVNEVFPAKTCTADESAE 60 
Dd	1 DAHSEVAHFRLDGBENFKALVLIAPAOYLQOCFFEDHVKLIVNEVFPAKTCTADESAE 60 
Qy	61 NCDKSLHTFGDKLCTVATFLRETGYGEMADCCAKOBERNECFLOHKDDNPNLPRIVRPEV 120 
Dd	61 NCDKSLHTFGDKLCTVATFLRETGYGEMADCCAKOBERNECFLOHKDDNPNLPRIVRPEV 120 
Qy	121 DVMCTAFHDNEETFLKKYLYETIARRHPHYFYAPPELLFAKRYKAATFTCCQAADKAACLLP 180 
Dd	121 DVMCTAFHDNEETFLKKYLYETIARRHPHYFYAPPELLFAKRYKAATFTCCQAADKAACLLP 180 
Qy	181 KIDELRDEGHASAKORLKCSLQKGERAFRAWAVALRSOPFPKAEFAVSKLIVTDLTk 240 
Dd	181 KIDELRDEGHASAKORLKCSLQKGERAFRAWAVALRSOPFPKAEFAVSKLIVTDLTk 240 
Qy	241 VHTECCGGDLLECADRDADLAKYICENODSISSKLKECERPILEKSHCIAFVENDEMPA 300 
Dd	241 VHTECCGGDLLECADRDADLAKYICENODSISSKLKECERPILEKSHCIAFVENDEMPA 300 
Qy	301 DLPLSAADFVESKDVCNKYAEDKVFLGMFLYEYARRHPDYSVVILLRLAKTYETTLEKC 360 
Dd	301 DLPLSAADFVESKDVCNKYAEDKVFLGMFLYEYARRHPDYSVVILLRLAKTYETTLEKC 360 
Qy	361 CAADPHECYAKVDFBFLPVEEPQNLIKONCELFEOJGEYKFONALLVRYTKYPQVST 420 
Dd	361 CAADPHECYAKVDFBFLPVEEPQNLIKONCELFEOJGEYKFONALLVRYTKYPQVST 420 
Qy	421 PTLVEVRNLGVSGCKCHPEAKRMPCADLYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480 
Dd	421 PTLVEVRNLGVSGCKCHPEAKRMPCADLYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480 
Qy	481 LVNRPPCFSALEYDETTPVKEFNAETTFTHADICTLSKERQIKKOTIALVEVLVKKPKAT 540 



```
Db 481 LVNRRPCFSALEVDYVYVKEAETFFHADICTLSEKERQIKQTALVELVHKPKAT 540
QY 541 KEQLKAYMDDFAAFVEKCKKADKTCFAEGRKLVAAQAALGL 585
Db 541 KEQLKAYMDDFAAFVEKCKKADKTCFAEGRKLVAAQAALGL 585

RESULT 15
ABG63321
ID ABG63321 standard; protein; 585 AA.
AC ABG63321;
XX
XX 27-AUG-2002 (first entry)
XX
XX Human serum albumin (HSA) protein.
XX
XX Albumin fusion protein; therapeutic protein X; human albumin; HA;
XX human serum albumin; HSA; cancer; reproductive disorder;
XX digestive disorder; immune disorder; endocrine disorder;
XX haematopoietic disorder; neural disorder; connective disorder;
XX cytostatic; antinfertility; antiinflammatory; antitumor;
XX immunomodulator; anti-HIV; antidiabetic; haemostatic; neurotropic;
XX neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
XX osteopathic; antiarthritic.
XX
XX Homo sapiens.
XX
XX OS
XX
XX WC20010717137-AL.
XX
XX 18-OCT-2001.
XX
XX 12-APR-2001; 2001WO-US11988.
XX
XX 12-APR-2000; 2000US-229358P.
XX
XX 23-APR-2000; 2000US-199384P.
XX
XX 21-DEC-2000; 2000US-256931P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Haseltine WA;
XX
XX WPI; 2002-010886/01.
XX
XX N-PSDB; ABK93280.
XX
XX New fusion protein for treating disease e.g. diabetes comprises an
XX albumin fused to a therapeutic protein -
XX
XX Claim 1; Fig 15; 2102pp; English.
XX
XX The present invention relates to albumin fusion proteins comprising a
XX therapeutic protein X and human albumin (HA), also known as human serum
XX albumin, HSA). The proteins are useful for treating a disease or
XX disorder that may be modulated by therapeutic protein X. The albumin
XX extends the shelf-life of protein X, and may increase its biological
XX in vitro/in vivo activity. The protein is useful for treating and
XX diagnosing disorders such as cancer, reproductive disorders, digestive
XX disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
XX (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
XX (e.g. diabetes), haematopoietic disorders, neural disorders
XX (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
XX encephalomyelitis, meningitis, schizophrenia), and connective disorders
XX (e.g. osteoporosis, arthritis). The present sequence represents HSA
XX (HA) protein.
XX
XX Sequence 585 AA;
XX
XX Query Match 100.0%; Score 3103; DB 23; Length 585;
XX Best Local Similarity 100.0%; Pred. No. 9.5e-255;
XX Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 DAHSEVAHFKDLSGENFKALVLAFAQLQCCPFEDHVKLVNVEYEFATCVADESAE 60
```

```
Db 1 DAHSEVAHFKDLSGENFKALVLAFAQLQCCPFEDHVKLVNVEYEFATCVADESAE 60
QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
Db 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKLYELIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
Db 121 DVMCTAFHDNEETFLKKLYELIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELDEGKASSAKORLKCSLQKGERAFKAWAVARLSQRPFAEVSKLIVTDLTK 240
Db 181 KLDELDEGKASSAKORLKCSLQKGERAFKAWAVARLSQRPFAEVSKLIVTDLTK 240
QY 241 VHTCCCHGDLLECADRADLAKYICENODSISSKLKECCCKPLLEKSHCTAEVENDENMPA 300
Db 241 VHTCCCHGDLLECADRADLAKYICENODSISSKLKECCCKPLLEKSHCTAEVENDENMPA 300
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYETARRHPDYSVILLRLAKTYETVILEK 360
Db 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYETARRHPDYSVILLRLAKTYETVILEK 360
QY 361 CAAADPHECTAKVDFEFPKLVPEEPQNLKQNCLEFQGLGEYKFQNALVRYTKVPOVST 420
Db 361 CAAADPHECTAKVDFEFPKLVPEEPQNLKQNCLEFQGLGEYKFQNALVRYTKVPOVST 420
QY 421 PTLVEVSRLNGYKSGCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRLNGYKSGCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYVYVKEAETFFHADICTLSEKERQIKQTALVELVHKPKAT 540
Db 481 LVNRRPCFSALEVDYVYVKEAETFFHADICTLSEKERQIKQTALVELVHKPKAT 540
QY 541 KEQLKAYMDDFAAFVEKCKKADKTCFAEGRKLVAAQAALGL 585
Db 541 KEQLKAYMDDFAAFVEKCKKADKTCFAEGRKLVAAQAALGL 585
```

Search completed: August 31, 2003, 16:38:41  
Job time : 88 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2003, 16:33:10 ; Search time 30 Seconds  
(without alignments)  
825.062 Million cell updates/sec

Title: US-09-833-041-18  
Perfect score: 3103  
Sequence: 1 DAKSEVAFKDLGEENFK.....TCFAEFGKLVAAQAALGL 585

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3103	100.0	585	1	US-08-153-799-14
2	3103	100.0	585	2	US-08-702-572-2
3	3103	100.0	585	3	US-08-769-746-2
4	3103	100.0	610	2	US-08-797-689-2
5	3103	100.0	783	1	US-08-256-938-2
6	3103	100.0	787	1	US-08-256-938-4
7	3103	100.0	787	2	US-08-797-689-16
8	3099	99.9	609	1	US-08-222-619-3
9	3099	99.9	609	1	US-08-433-037-4
10	3099	99.9	609	4	US-08-897-956A-2
11	3099	99.9	609	5	PCT-US95-04075-3
12	3095	99.7	978	4	US-08-897-956A-3
13	3093	99.7	585	1	US-08-448-196A-3
14	3093	99.7	585	2	US-08-984-176-1
15	2458.5	79.2	583	1	US-08-448-196A-5
16	2450.5	79.0	583	1	US-08-448-196A-4
17	2432.5	78.4	583	1	US-08-448-196A-6
18	2426	78.2	584	1	US-08-448-196A-7
19	2389	77.0	582	1	US-08-134-638-1
20	1249.5	40.3	609	1	US-08-222-619-4
21	1249.5	40.3	609	5	PCT-US95-04075-4
22	1206.5	38.9	590	2	US-08-377-309-2
23	1206.5	38.9	590	3	US-09-186-723-2
24	1206.5	38.9	590	4	US-08-505-012-5
25	1206.5	38.9	590	4	US-09-186-949A-3
26	1206.5	38.9	590	4	US-08-758-757-2
27	1206.5	38.9	590	5	PCT-US96-00996-5

28 1206.5 38.9 609 4 US-09-186-949A-2  
29 1164.5 37.5 579 1 US-08-448-196A-8  
30 1055 34.0 599 1 US-08-222-619-2  
31 1055 34.0 599 3 US-08-221-767-24  
32 1055 34.0 599 5 PCT-US95-04075-2  
33 926 29.8 393 2 US-08-377-309-7  
34 926 29.8 393 3 US-09-186-723-7  
35 926 29.8 393 4 US-08-503-012-10  
36 926 29.8 393 4 US-09-186-949A-8  
37 926 29.8 393 4 US-08-758-757-7  
38 926 29.8 393 5 PCT-US96-00996-10  
39 777 25.0 324 4 US-08-505-012-12  
40 777 25.0 324 5 PCT-US96-00996-12  
41 777 25.0 325 2 US-08-377-309-8  
42 777 25.0 325 3 US-09-186-723-8  
43 777 25.0 325 4 US-08-505-012-11  
44 777 25.0 325 4 US-09-186-949A-9  
45 777 25.0 325 4 US-08-758-757-8

ALIGNMENTS

RESULT 1  
US-08-153-799-14  
; Sequence 14, Application US/08153799  
; Patent No. 5766883  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J  
; APPLICANT: Goodey, Andrew R  
; TITLE OF INVENTION: Polypeptides  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: R Hain Swope, BOC Health Care Inc  
; STREET: 100 Mountain Avenue  
; CITY: Murray Hill  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07974  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/153,799  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/847975  
; FILING DATE: 06-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8909916.2  
; FILING DATE: 29-APR-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB90/00650  
; FILING DATE: 26-APR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/775952  
; FILING DATE: 29-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Swope, R Hain  
; REGISTRATION NUMBER: 24864  
; REFERENCE/DOCKET NUMBER: 92H832  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 665 2400  
; TELEFAX: (908) 771 6159  
; TELEX: 219484  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: Region  
 ; LOCATION: 369..419  
 ; OTHER INFORMATION: /note= "Alternative C-termini of  
 ; OTHER INFORMATION: HSA(1-n)"  
 ; FEATURE:  
 ; NAME/KEY: Region  
 ; LOCATION: 1..585  
 ; OTHER INFORMATION: /note= "Amino acid sequence of  
 ; OTHER INFORMATION: natural HSA"  
 ; US-08-153-799-14

Query Match 100.0%; Score 3103; DB 1; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 9.3e-287;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DAHSEVAHFRKDLGSENFKALVLI	FAQYLQCCPFEDHVKLVNVE	TEFAKTCVAD	ESAE	60
Db	1	DAHSEVAHFRKDLGSENFKALVLI	FAQYLQCCPFEDHVKLVNVE	TEFAKTCVAD	ESAE	60
QY	61	NCDKSLHTLPGDKLCTVATL	RETYGEMADCCAKOE	PERNECFLOHKDDNP	NPLVLRPEV	120
Db	61	NCDKSLHTLPGDKLCTVATL	RETYGEMADCCAKOE	PERNECFLOHKDDNP	NPLVLRPEV	120
QY	121	DVMCTAFHDNEETFLKKY	LIETARRHPYFYAP	ELLFPKRYKAAPTECCQ	AAKACLLP	180
Db	121	DVMCTAFHDNEETFLKKY	LIETARRHPYFYAP	ELLFPKRYKAAPTECCQ	AAKACLLP	180
QY	181	KLDELDEGKASSAKQRLK	CASLQKFGERA	FKAWAVARLSQRP	PKAFAEYSKLVTDLTK	240
Db	181	KLDELDEGKASSAKQRLK	CASLQKFGERA	FKAWAVARLSQRP	PKAFAEYSKLVTDLTK	240
QY	241	VHTECCGGDILLECADD	RADLAKYICENQDS	ISSKLEKCEKPLLEK	SHCIAEVENDEMPA	300
Db	241	VHTECCGGDILLECADD	RADLAKYICENQDS	ISSKLEKCEKPLLEK	SHCIAEVENDEMPA	300
QY	301	DLPSLAADPFVESKD	VCKNYAEAKDVFLG	MFLEYARRHPDYSV	LLRLAKTYETTLTK	360
Db	301	DLPSLAADPFVESKD	VCKNYAEAKDVFLG	MFLEYARRHPDYSV	LLRLAKTYETTLTK	360
QY	361	CAAADPHECYAKVDFE	KPLVEEPONL	KQNCSELFQLG	EYKFNALLVRYTKKVPQVST	420
Db	361	CAAADPHECYAKVDFE	KPLVEEPONL	KQNCSELFQLG	EYKFNALLVRYTKKVPQVST	420
QY	421	PTLVEVSRNLGKVGSK	CKKHPKEMPCAED	YLSVVLNQLCVL	HEKTPVSDRVTKCTES	480
Db	421	PTLVEVSRNLGKVGSK	CKKHPKEMPCAED	YLSVVLNQLCVL	HEKTPVSDRVTKCTES	480
QY	481	LVNRRPCFSALEVDE	TYVPKEFNAETFF	HADICTLSEKERO	IKKQTAALVELVHKPKAT	540
Db	481	LVNRRPCFSALEVDE	TYVPKEFNAETFF	HADICTLSEKERO	IKKQTAALVELVHKPKAT	540
QY	541	KEQLKAYMDPFAAF	VEKCKADKCTCF	FAEEGKKLVAASQA	ALGL	585
Db	541	KEQLKAYMDPFAAF	VEKCKADKCTCF	FAEEGKKLVAASQA	ALGL	585

RESULT 2  
 US-08-702-572-2  
 ; Sequence 2, Application US/08702572  
 ; Patent No. 5965386  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kerry-Williams, Sean M  
 ; APPLICANT: Gilbert, Sarah C  
 ; TITLE OF INVENTION: Yeast Strains and Modified Albumins  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: Centeon L.L.C.  
 ; STREET: 1020 First Avenue

; CITY: King of Prussia  
 ; STATE: Pennsylvania  
 ; COUNTRY: USA  
 ; ZIP: 19406-1310  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: MS-DOS  
 ; SOFTWARE: Microsoft Word 6.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/702,572  
 ; FILING DATE: 11-NOV-1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: WO 95/23857  
 ; FILING DATE: 1-MAR-1995  
 ; APPLICATION NUMBER: GB 9404270.2  
 ; FILING DATE: 5-MAR-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Naomi Biswas  
 ; REGISTRATION NUMBER: 38,384  
 ; REFERENCE/DOCKET NUMBER: CE0114 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 610/878/4294  
 ; TELEFAX: 610/878/4221  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 585 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-702-572-2

Query Match 100.0%; Score 3103; DB 2; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 9.3e-287;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DAHSEVAHFRKDLGSENFKALVLI	FAQYLQCCPFEDHVKLVNVE	TEFAKTCVAD	ESAE	60
Db	1	DAHSEVAHFRKDLGSENFKALVLI	FAQYLQCCPFEDHVKLVNVE	TEFAKTCVAD	ESAE	60
QY	61	NCDKSLHTLPGDKLCTVATL	RETYGEMADCCAKOE	PERNECFLOHKDDNP	NPLVLRPEV	120
Db	61	NCDKSLHTLPGDKLCTVATL	RETYGEMADCCAKOE	PERNECFLOHKDDNP	NPLVLRPEV	120
QY	121	DVMCTAFHDNEETFLKKY	LIETARRHPYFYAP	ELLFPKRYKAAPTECCQ	AAKACLLP	180
Db	121	DVMCTAFHDNEETFLKKY	LIETARRHPYFYAP	ELLFPKRYKAAPTECCQ	AAKACLLP	180
QY	181	KLDELDEGKASSAKQRLK	CASLQKFGERA	FKAWAVARLSQRP	PKAFAEYSKLVTDLTK	240
Db	181	KLDELDEGKASSAKQRLK	CASLQKFGERA	FKAWAVARLSQRP	PKAFAEYSKLVTDLTK	240
QY	241	VHTECCGGDILLECADD	RADLAKYICENQDS	ISSKLEKCEKPLLEK	SHCIAEVENDEMPA	300
Db	241	VHTECCGGDILLECADD	RADLAKYICENQDS	ISSKLEKCEKPLLEK	SHCIAEVENDEMPA	300
QY	301	DLPSLAADPFVESKD	VCKNYAEAKDVFLG	MFLEYARRHPDYSV	LLRLAKTYETTLTK	360
Db	301	DLPSLAADPFVESKD	VCKNYAEAKDVFLG	MFLEYARRHPDYSV	LLRLAKTYETTLTK	360
QY	361	CAAADPHECYAKVDFE	KPLVEEPONL	KQNCSELFQLG	EYKFNALLVRYTKKVPQVST	420
Db	361	CAAADPHECYAKVDFE	KPLVEEPONL	KQNCSELFQLG	EYKFNALLVRYTKKVPQVST	420
QY	421	PTLVEVSRNLGKVGSK	CKKHPKEMPCAED	YLSVVLNQLCVL	HEKTPVSDRVTKCTES	480
Db	421	PTLVEVSRNLGKVGSK	CKKHPKEMPCAED	YLSVVLNQLCVL	HEKTPVSDRVTKCTES	480
QY	481	LVNRRPCFSALEVDE	TYVPKEFNAETFF	HADICTLSEKERO	IKKQTAALVELVHKPKAT	540
Db	481	LVNRRPCFSALEVDE	TYVPKEFNAETFF	HADICTLSEKERO	IKKQTAALVELVHKPKAT	540

QY 541 KEOLKAVMDFAAFVEKCKKADKCTCFABEGKKLVAASQAALGL 585  
Db 541 KEOLKAVMDFAAFVEKCKKADKCTCFABEGKKLVAASQAALGL 585

## RESULT 3

US-08-769-746-2  
; Sequence 2, Application US/08759746  
; Patent No. 6274305  
; GENERAL INFORMATION:  
; APPLICANT: Sonnschein, Carlos  
; APPLICANT: Soto, Ana M.  
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/769,746  
; FILING DATE: 19-DEC-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: MBRI-02584  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 2:  
; LENGTH: 585 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-769-746-2

Query Match 100.0%; Score 3103; DB 3; Length 585;  
Best Local Similarity 100.0%; Pred. No. 9.3e-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHFKDLGEENFKALVLIAPQYLQCCPPEDHVKLVNVEFTAKTCVADDSAE 60  
Db 1 DAHKEVAHFKDLGEENFKALVLIAPQYLQCCPPEDHVKLVNVEFTAKTCVADDSAE 60  
QY 61 NCKDSLHTFLGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPLRVEPV 120  
Db 61 NCKDSLHTFLGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPLRVEPV 120  
QY 121 DVMCTAFHDNEETFLKKLYETARRHPYFYAPPELLFAKRYKAATECCQADKAACLLP 180  
Db 121 DVMCTAFHDNEETFLKKLYETARRHPYFYAPPELLFAKRYKAATECCQADKAACLLP 180  
QY 181 KLDELDRGKASSAKQRIKACSIQKFGERAFKANAVARLSQRPKAEFAEYSKLVITDITK 240  
Db 181 KLDELDRGKASSAKQRIKACSIQKFGERAFKANAVARLSQRPKAEFAEYSKLVITDITK 240  
QY 241 VITECHGDLLLECADDRADIAYICENODSISSKKECCCKPLLEKSHCIAFVENDMPA 300  
Db 241 VITECHGDLLLECADDRADIAYICENODSISSKKECCCKPLLEKSHCIAFVENDMPA 300  
QY 301 DLPSTAADFVSKDVCKNVAERKDVFLGMFLYEVARRHPDYSVVLLLRKATYETTLK 360  
Db 301 DLPSTAADFVSKDVCKNVAERKDVFLGMFLYEVARRHPDYSVVLLLRKATYETTLK 360

QY 361 CAAADPHCYAKVDEEFKPLVEBPQNLKQNCSELFQOLGEYKFNALLVRYTKKYPQVST 420  
Db 361 CAAADPHCYAKVDEEFKPLVEBPQNLKQNCSELFQOLGEYKFNALLVRYTKKYPQVST 420  
QY 421 PTLVEVSRNLKGVSKCKKHPKAMPKADYLSVNLNQLCVLHEKTPYSDRVTKCCTES 480  
Db 421 PTLVEVSRNLKGVSKCKKHPKAMPKADYLSVNLNQLCVLHEKTPYSDRVTKCCTES 480  
QY 481 LVNRPCFSALEVDYETVPKSENAETFTFHADICTLSEKEROIKKOTALVELVKKHKPKAT 540  
Db 481 LVNRPCFSALEVDYETVPKSENAETFTFHADICTLSEKEROIKKOTALVELVKKHKPKAT 540  
QY 541 KEOLKAVMDFAAFVEKCKKADKCTCFABEGKKLVAASQAALGL 585  
Db 541 KEOLKAVMDFAAFVEKCKKADKCTCFABEGKKLVAASQAALGL 585

## RESULT 4

US-08-797-689-2  
; Sequence 2, Application US/08797689  
; Patent No. 5876969  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; APPLICANT: Fournier, Alain  
; APPLICANT: Guillon, Jean-Dominique  
; APPLICANT: Jung, Gerard  
; APPLICANT: Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3043  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.1 (Patentin)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/797,689  
; FILING DATE: 31-JAN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/256,927  
; FILING DATE: 28-JUL-1994  
; APPLICATION NUMBER: FR 92/01064  
; FILING DATE: 31-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR93/00085  
; FILING DATE: 28-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith Ph.D., Julie K.  
; REGISTRATION NUMBER: P-38,619  
; REFERENCE/DOCKET NUMBER: ST92006-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3839  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-797-689-2

Query Match 100.0%; Score 3103; DB 2; Length 610;  
Best Local Similarity 100.0%; Pred. No. 9.8e-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGEENFKALVLIATFAQYLOQCPFDHVKLVNVEYFAKTCVADSAE 60  
Db 25 DAHSEVAHRFKDGEENFKALVLIATFAQYLOQCPFDHVKLVNVEYFAKTCVADSAE 84  
QY 61 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKOEPERNECFLOHKDDNPMLPLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKOEPERNECFLOHKDDNPMLPLVRPEV 144  
QY 121 DVMCTAFHDNEETFLAKYIETARHPYFYAPPELLFAKRYKAATFECQQAADKAACLLP 180  
Db 145 DVMCTAFHDNEETFLAKYIETARHPYFYAPPELLFAKRYKAATFECQQAADKAACLLP 204  
QY 181 KLDELREBGRASSAKQRLKASLOKFGERAFKAWAVARLSQRPKAEFAEYSKLVTDLT 240  
Db 205 KLDELREBGRASSAKQRLKASLOKFGERAFKAWAVARLSQRPKAEFAEYSKLVTDLT 264  
QY 241 VTECCGHDLLCADDADRLAKYICENODSISSKLKCCCKPILLESKSHCIAEVENDEMPA 300  
Db 265 VTECCGHDLLCADDADRLAKYICENODSISSKLKCCCKPILLESKSHCIAEVENDEMPA 324  
QY 301 DPLSLAADFVESKDVCKNYAEAKDVLGMLFYETARHPDY SVLLRLAKTYETLEK 360  
Db 325 DPLSLAADFVESKDVCKNYAEAKDVLGMLFYETARHPDY SVLLRLAKTYETLEK 384  
QY 361 CAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNQALVRYTKKVPQVST 420  
Db 385 CAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNQALVRYTKKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVNLQNLVLEHPTVSDRVTKCCTES 480  
Db 445 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVNLQNLVLEHPTVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDYETVYKPEFNAETFFHADICTLSEKEROIKQTALVELVKKPKAT 540  
Db 505 LVNRRPCFSALEVDYETVYKPEFNAETFFHADICTLSEKEROIKQTALVELVKKPKAT 564  
QY 541 KEQLKAVMDFFAAFEVCKCKKADDKETCFABEGKKLVAASQAALGL 585  
Db 565 KEQLKAVMDFFAAFEVCKCKKADDKETCFABEGKKLVAASQAALGL 609

## RESULT 5

US-08-256-938-2  
; Sequence 2, Application US/08256938  
; Patent No. 5665863  
; GENERAL INFORMATION:  
; APPLICANT: Yeh, Patrice  
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE  
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.0 (PatentIn)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/256,938  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 92/01065  
; FILING DATE: 31-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goodman, Rosanne

; REGISTRATION NUMBER: 32,534  
; REFERENCE/DOCKET NUMBER: ST92007-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3817  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 783 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-256-938-2

Query Match 100.0%; Score 3103; DB 1; Length 783;  
Best Local Similarity 100.0%; Pred. No. 1.4e-286;  
Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHRFKDGEENFKALVLIATFAQYLOQCPFDHVKLVNVEYFAKTCVADSAE 60  
Db 25 DAHSEVAHRFKDGEENFKALVLIATFAQYLOQCPFDHVKLVNVEYFAKTCVADSAE 84  
QY 61 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKOEPERNECFLOHKDDNPMLPLVRPEV 120  
Db 85 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKOEPERNECFLOHKDDNPMLPLVRPEV 144  
QY 121 DVMCTAFHDNEETFLAKYIETARHPYFYAPPELLFAKRYKAATFECQQAADKAACLLP 180  
Db 145 DVMCTAFHDNEETFLAKYIETARHPYFYAPPELLFAKRYKAATFECQQAADKAACLLP 204  
QY 181 KLDELREBGRASSAKQRLKASLOKFGERAFKAWAVARLSQRPKAEFAEYSKLVTDLT 240  
Db 205 KLDELREBGRASSAKQRLKASLOKFGERAFKAWAVARLSQRPKAEFAEYSKLVTDLT 264  
QY 241 VTECCGHDLLCADDADRLAKYICENODSISSKLKCCCKPILLESKSHCIAEVENDEMPA 300  
Db 265 VTECCGHDLLCADDADRLAKYICENODSISSKLKCCCKPILLESKSHCIAEVENDEMPA 324  
QY 301 DPLSLAADFVESKDVCKNYAEAKDVLGMLFYETARHPDY SVLLRLAKTYETLEK 360  
Db 325 DPLSLAADFVESKDVCKNYAEAKDVLGMLFYETARHPDY SVLLRLAKTYETLEK 384  
QY 361 CAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNQALVRYTKKVPQVST 420  
Db 385 CAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNQALVRYTKKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVNLQNLVLEHPTVSDRVTKCCTES 480  
Db 445 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVNLQNLVLEHPTVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDYETVYKPEFNAETFFHADICTLSEKEROIKQTALVELVKKPKAT 540  
Db 505 LVNRRPCFSALEVDYETVYKPEFNAETFFHADICTLSEKEROIKQTALVELVKKPKAT 564  
QY 541 KEQLKAVMDFFAAFEVCKCKKADDKETCFABEGKKLVAASQAALGL 585  
Db 565 KEQLKAVMDFFAAFEVCKCKKADDKETCFABEGKKLVAASQAALGL 609

## RESULT 6

US-08-256-938-4  
; Sequence 4, Application US/08256938  
; Patent No. 5665863  
; GENERAL INFORMATION:  
; APPLICANT: Yeh, Patrice  
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE  
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA

COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.0 (Patentin)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,938  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 92/01065  
FILING DATE: 31-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Goodman, Rosanne  
REGISTRATION NUMBER: 32,534  
REFERENCE/DOCKET NUMBER: ST92007-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3817  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 787 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-256-938-4

Query Match 100.0%; Score 3103; DB 1; Length 787;  
Best Local Similarity 100.0%; Pred. No. 1.4e-286;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDGLGENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADSEAE 60  
DB 203 DAHSEVAHFKDGLGENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADSEAE 262

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVPEV 120  
DB 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVPEV 322

QY 121 DVMCTAFHDMNEETFLKKYLYETARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
DB 323 DVMCTAFHDMNEETFLKKYLYETARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 382

QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWARLSORPPKAEFAEVSPLYTDLTK 240  
DB 383 KLDELDEGKASSAKORLKASLOKFGERAFKAWARLSORPPKAEFAEVSPLYTDLTK 442

QY 241 VTECHGDLLECAADDRADIACYICENQDSISSKLECECKPILLESKSHCIAEVNDMPA 300  
DB 443 VTECHGDLLECAADDRADIACYICENQDSISSKLECECKPILLESKSHCIAEVNDMPA 502

QY 301 DLPSLAADFVSKDVCKNTAEAKDVLGMFLYETARRHPDYSVVLLRLAKYETTLTK 360  
DB 503 DLPSLAADFVSKDVCKNTAEAKDVLGMFLYETARRHPDYSVVLLRLAKYETTLTK 562

QY 361 CAADPHCYAKVDEKPLVEPQNLKONCELFQOLGYKFNALIVRTTKVPOVST 420  
DB 563 CAADPHCYAKVDEKPLVEPQNLKONCELFQOLGYKFNALIVRTTKVPOVST 622

QY 421 PTLVEVSRNLKGVSKCKHPEAKRMPCAEYLSVLNQLCVLHEKTPVSDRVTKCTES 480  
DB 623 PTLVEVSRNLKGVSKCKHPEAKRMPCAEYLSVLNQLCVLHEKTPVSDRVTKCTES 682

QY 481 LVNRPCTSALEVDVTPKEFNAETFFHADICTLSEKROIKKQTALVELVKKPKAT 540  
DB 683 LVNRPCTSALEVDVTPKEFNAETFFHADICTLSEKROIKKQTALVELVKKPKAT 742

QY 541 KEOLKAVMDFAAEVFECKCKADDKETCPAEFGKKVNAASQAALGL 585  
DB 743 KEOLKAVMDFAAEVFECKCKADDKETCPAEFGKKVNAASQAALGL 787

RESULT 7  
US-08-797-689-16  
Sequence 16, Application US/08797689  
Patent No. 5876969  
GENERAL INFORMATION:  
APPLICANT: Fleer, Reinhard  
APPLICANT: Fournier, Alain  
APPLICANT: Guittou, Jean-Dominique  
APPLICANT: Jung, Gerard  
APPLICANT: Yeh, Patrice  
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESS: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3043  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (Patentin)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 787 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-797-689-16

Query Match 100.0%; Score 3103; DB 2; Length 787;  
Best Local Similarity 100.0%; Pred. No. 1.4e-286;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDGLGENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADSEAE 60  
DB 203 DAHSEVAHFKDGLGENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADSEAE 262

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVPEV 120  
DB 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVPEV 322

QY 121 DVMCTAFHDMNEETFLKKYLYETARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
DB 323 DVMCTAFHDMNEETFLKKYLYETARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 382

QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWARLSORPPKAEFAEVSPLYTDLTK 240  
DB 383 KLDELDEGKASSAKORLKASLOKFGERAFKAWARLSORPPKAEFAEVSPLYTDLTK 442

QY 241 VHTCCGDLLECGADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB 443 VHTCCGDLLECGADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 502  
QY 301 DLPSLAADFVESKDVCKNAEAKDVFLGMFLYFYARRHPDYSVWLLRLAKTYETTTLEK 360  
DB 503 DLPSLAADFVESKDVCKNAEAKDVFLGMFLYFYARRHPDYSVWLLRLAKTYETTTLEK 562  
QY 361 CAADPHCEYAKVDFEYKPLVEEPONLIKONCELFQELGEYKFNALLVRYTKKVPQVST 420  
DB 563 CAADPHCEYAKVDFEYKPLVEEPONLIKONCELFQELGEYKFNALLVRYTKKVPQVST 522  
QY 421 PTLVEVSRNLGKVGSKCKHPKMPCAEDYLSVNLQCVLHKTPTVSDRVTKCCTES 480  
DB 623 PTLVEVSRNLGKVGSKCKHPKMPCAEDYLSVNLQCVLHKTPTVSDRVTKCCTES 682  
QY 481 LYNRRPCFSALEVDEYVVPKEFNAETFFHADICTLSEKEROIKKOTALVELVKKPKAT 540  
DB 683 LYNRRPCFSALEVDEYVVPKEFNAETFFHADICTLSEKEROIKKOTALVELVKKPKAT 742  
QY 541 KEOLKAVMDFFAAVFEKCKADDDKTCFAEGKKLVAASQAALGL 585  
DB 743 KEOLKAVMDFFAAVFEKCKADDDKTCFAEGKKLVAASQAALGL 787

## RESULT 8

US-08-222-619-3  
; Sequence 3, Application US/08222619  
; Patent No. 5652352  
; GENERAL INFORMATION:  
; APPLICANT: Lichenstein, Henri  
; APPLICANT: Lyons, David  
; APPLICANT: Wurfel, Mark  
; APPLICANT: Wright, Samuel  
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like  
; TITLE OF INVENTION: Protein  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amgen Center, Patent Operations/RRC  
; STREET: 1840 DeHavilland Drive  
; CITY: Thousand Oaks  
; STATE: California  
; COUNTRY: U.S.  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: US/08/222,619  
; CLASSIFICATION: 435  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 609 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-222-619-3

Query Match 99.9%; Score 3099; DB 1; Length 609;  
Best Local Similarity 99.8%; Pred. No. 2.4e-286;  
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVARRKFDLGEENFKALVLIAPQYLQCCPFEDHVKLVNTEYFAKTCVADESAAE 60  
DB 25 DAHSEVARRKFDLGEENFKALVLIAPQYLQCCPFEDHVKLVNTEYFAKTCVADESAAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPNLPLRVREV 120

DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPNLPLRVREV 144  
QY 121 DVMCTAFHONETFTKKVLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCQAAAKACLIP 180  
DB 145 DVMCTAFHONETFTKKVLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCQAAAKACLIP 204  
QY 181 KIDELRDEGKASAKORLKASLOKFGERAFKAWAVARLSORFFKAFVAVSKLVTDLTG 240  
DB 205 KIDELRDEGKASAKORLKASLOKFGERAFKAWAVARLSORFFKAFVAVSKLVTDLTG 264  
QY 241 VHTCCGDLLECGADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB 265 VHTCCGDLLECGADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNAEAKDVFLGMFLYFYARRHPDYSVWLLRLAKTYETTTLEK 360  
DB 325 DLPSLAADFVESKDVCKNAEAKDVFLGMFLYFYARRHPDYSVWLLRLAKTYETTTLEK 384  
QY 361 CAADPHCEYAKVDFEYKPLVEEPONLIKONCELFQELGEYKFNALLVRYTKKVPQVST 420  
DB 385 CAADPHCEYAKVDFEYKPLVEEPONLIKONCELFQELGEYKFNALLVRYTKKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKHPKMPCAEDYLSVNLQCVLHKTPTVSDRVTKCCTES 480  
DB 445 PTLVEVSRNLGKVGSKCKHPKMPCAEDYLSVNLQCVLHKTPTVSDRVTKCCTES 504  
QY 481 LYNRRPCFSALEVDEYVVPKEFNAETFFHADICTLSEKEROIKKOTALVELVKKPKAT 540  
DB 505 LYNRRPCFSALEVDEYVVPKEFNAETFFHADICTLSEKEROIKKOTALVELVKKPKAT 564  
QY 541 KEOLKAVMDFFAAVFEKCKADDDKTCFAEGKKLVAASQAALGL 585  
DB 565 KEOLKAVMDFFAAVFEKCKADDDKTCFAEGKKLVAASQAALGL 609

## RESULT 9

US-08-433-037-4  
; Sequence 4, Application US/08433037  
; Patent No. 5707828  
; GENERAL INFORMATION:  
; APPLICANT: Sreekrishna, Kotikanyadan  
; APPLICANT: Barr, Kathryn A.  
; APPLICANT: Bristerley, Russell A.  
; APPLICANT: Thrill, Gregory P.  
; APPLICANT: Tschopp, Juerg F.  
; TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN  
; TITLE OF INVENTION: PICHIA PASTORIS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 11530-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/433,037  
; FILING DATE: 03-MAY-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 9108Z  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:  
LENGTH: 609 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-433-037-4

Query Match 99.9%; Score 3099; DB 1; Length 609;  
Best local similarity 99.8%; Pred. No. 2.4e-286;  
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHKSVAHREFKDLGEENFKALVLIAPAOYLQOCPPFEDHVKLVNVEYTFEFAKTCVADESA 60  
DB 25 DAHKSVAHREFKDLGEENFKALVLIAPAOYLQOCPPFEDHVKLVNVEYTFEFAKTCVADESA 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDDNPNLRLVREPV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDDNPNLRLVREPV 144  
QY 121 DVMCTAFHNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAFTCCQAADKAACLLP 180  
DB 145 DVMCTAFHNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAFTCCQAADKAACLLP 204  
QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRFPAFAEVSCLVDTLTK 240  
DB 205 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRFPAFAEVSCLVDTLTK 264  
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB 265 VHTCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKOVCKNYAEAKDVFGLMFLEYIARRHPDYSVLLRLAKTYETTLK 360  
DB 325 DLPSLAADFVESKOVCKNYAEAKDVFGLMFLEYIARRHPDYSVLLRLAKTYETTLK 384  
QY 361 CAAADPHECYAKVDFEFLVPEEPONLIKONCELFQGEYKFQNALLVYTKVPQVST 420  
DB 385 CAAADPHECYAKVDFEFLVPEEPONLIKONCELFQGEYKFQNALLVYTKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480  
DB 445 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 504  
QY 481 LVNRRPCFSALEVDYTYVPKEFNAETFFHADICTLSEKEROIKQATLVELVKKPKAT 540  
DB 505 LVNRRPCFSALEVDYTYVPKEFNAETFFHADICTLSEKEROIKQATLVELVKKPKAT 564  
QY 541 KEQLKAVMDDFAAFEVKCKCKADKCTCFABEGKKLVAASQAALGL 585  
DB 565 KEQLKAVMDDFAAFEVKCKCKADKCTCFABEGKKLVAASQAALGL 609

RESULT 10  
US-08-897-956A-2  
Sequence 2, Application US/08897956A  
Patent No. 6423512  
GENERAL INFORMATION:  
APPLICANT: Mary Ellen Digan  
APPLICANT: Philip Lake  
APPLICANT: Hermann Gram  
TITLE OF INVENTION: Fusion Polypeptides  
FILE REFERENCE: 600-7244/CPA  
CURRENT APPLICATION NUMBER: US/08/897,956A  
CURRENT FILING DATE: 1997-07-21  
PRIOR APPLICATION NUMBER: 60/022,689  
PRIOR FILING DATE: 1996-07-26  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 609  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-08-897-956A-2

Query Match 99.9%; Score 3099; DB 4; Length 609;  
Best local similarity 99.8%; Pred. No. 2.4e-286;  
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHREFKDLGEENFKALVLIAPAOYLQOCPPFEDHVKLVNVEYTFEFAKTCVADESA 60  
DB 25 DAHKSVAHREFKDLGEENFKALVLIAPAOYLQOCPPFEDHVKLVNVEYTFEFAKTCVADESA 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDDNPNLRLVREPV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDDNPNLRLVREPV 144  
QY 121 DVMCTAFHNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAFTCCQAADKAACLLP 180  
DB 145 DVMCTAFHNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAFTCCQAADKAACLLP 204  
QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRFPAFAEVSCLVDTLTK 240  
DB 205 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRFPAFAEVSCLVDTLTK 264  
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB 265 VHTCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKOVCKNYAEAKDVFGLMFLEYIARRHPDYSVLLRLAKTYETTLK 360  
DB 325 DLPSLAADFVESKOVCKNYAEAKDVFGLMFLEYIARRHPDYSVLLRLAKTYETTLK 384  
QY 361 CAAADPHECYAKVDFEFLVPEEPONLIKONCELFQGEYKFQNALLVYTKVPQVST 420  
DB 385 CAAADPHECYAKVDFEFLVPEEPONLIKONCELFQGEYKFQNALLVYTKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480  
DB 445 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 504  
QY 481 LVNRRPCFSALEVDYTYVPKEFNAETFFHADICTLSEKEROIKQATLVELVKKPKAT 540  
DB 505 LVNRRPCFSALEVDYTYVPKEFNAETFFHADICTLSEKEROIKQATLVELVKKPKAT 564  
QY 541 KEQLKAVMDDFAAFEVKCKCKADKCTCFABEGKKLVAASQAALGL 585  
DB 565 KEQLKAVMDDFAAFEVKCKCKADKCTCFABEGKKLVAASQAALGL 609

RESULT 11  
PCT-US95-04075-3  
Sequence 3, Application PC/TUS9504075  
GENERAL INFORMATION:  
APPLICANT: AMGEN INC.  
TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like  
TITLE OF INVENTION: Protein  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Center, Patent Operations/RR  
STREET: 1840 DeHavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: U.S.  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04075  
FILING DATE:  
CLASSIFICATION:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 609 amino acids



; TYPE: amino acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 PCT-US95-04075-3

Query Match 99.9%; Score 3099; DB 5; Length 609;  
 Best Local Similarity 99.8%; Pred. No. 2.4e-286;  
 Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYLQCCPPEDHVKLVNVEVTEFAKTCVADESAE 60  
 Db 25 DAHSEVAHFRKDLGEENFKALVLIAPAQYLQCCPPEDHVKLVNVEVTEFAKTCVADESAE 84  
 QY 61 NCKSLHTLFGDKLCIVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLVRPEV 120  
 Db 85 NCKSLHTLFGDKLCIVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLVRPEV 144  
 QY 121 DVMCTAFHDNEETFLKKLYIETARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
 Db 145 DVMCTAFHDNEETFLKKLYIETARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 204  
 QY 181 KLDELREGKASSAKORLKCSLQKFGERAFKAWAVARLSQRPKAEFAEYVKLVITDLTK 240  
 Db 205 KLDELREGKASSAKORLKCSLQKFGERAFKAWAVARLSQRPKAEFAEYVKLVITDLTK 264  
 QY 241 VHTCCCHGDLLECADRADLAKYICENODSISSKLKECCCKPFLLEKSHCIAEVENDEMPA 300  
 Db 265 VHTCCCHGDLLECADRADLAKYICENODSISSKLKECCCKPFLLEKSHCIAEVENDEMPA 324  
 QY 301 DLPSLAADFVESKDVCKNTAEAKOVFLGMFLYETARRHPDYSVVLLRLAKTYETTTLEK 360  
 Db 325 DLPSLAADFVESKDVCKNTAEAKOVFLGMFLYETARRHPDYSVVLLRLAKTYETTTLEK 384  
 QY 361 CAADPEHCYAKVDFEDEFKPLVEEPONLIKONCELFQOLGEYFQNALVRYTKKVPQVST 420  
 Db 385 CAADPEHCYAKVDFEDEFKPLVEEPONLIKONCELFQOLGEYFQNALVRYTKKVPQVST 444  
 QY 421 PTLVEVSRLNGKVGSKCKKHPKAMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
 Db 445 PTLVEVSRLNGKVGSKCKKHPKAMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504  
 QY 481 LVNRRPCFSALEVDETYVPEKFNATFTFHADICTLSEKERQIKKQTALVELVHKHPKAT 540  
 Db 505 LVNRRPCFSALEVDETYVPEKFNATFTFHADICTLSEKERQIKKQTALVELVHKHPKAT 564  
 QY 541 KEOLKAYMDDFAAFVEKCKCKADDDKTCFAEBGKKLVAASQAALGL 585  
 Db 565 KEOLKAYMDDFAAFVEKCKCKADDDKTCFAEBGKKLVAASQAALGL 609

## RESULT 12

US-08-897-956A-3  
 ; Sequence 3, Application US/08897956A  
 ; Patent No. 6423512  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mary Ellen Digan  
 ; APPLICANT: Philip Lake  
 ; APPLICANT: Hermann Gram  
 ; TITLE OF INVENTION: Fusion Polypeptides  
 ; FILE REFERENCE: 600-7244/CPA  
 ; CURRENT APPLICATION NUMBER: US/08/897,956A  
 ; CURRENT FILING DATE: 1997-07-21  
 ; PRIOR APPLICATION NUMBER: 60/022,689  
 ; PRIOR FILING DATE: 1996-07-26  
 ; NUMBER OF SEQ ID NOS: 38  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 978  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Fusion polypeptide

## US-08-897-956A-3

Query Match 99.7%; Score 3095; DB 4; Length 978;  
 Best Local Similarity 99.8%; Pred. No. 1.1e-285;  
 Matches 583; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYLQCCPPEDHVKLVNVEVTEFAKTCVADESAE 60  
 Db 212 DAHSEVAHFRKDLGEENFKALVLIAPAQYLQCCPPEDHVKLVNVEVTEFAKTCVADESAE 271  
 QY 61 NCKSLHTLFGDKLCIVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLVRPEV 120  
 Db 272 NCKSLHTLFGDKLCIVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLVRPEV 331  
 QY 121 DVMCTAFHDNEETFLKKLYIETARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
 Db 332 DVMCTAFHDNEETFLKKLYIETARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 391  
 QY 181 KLDELREGKASSAKORLKCSLQKFGERAFKAWAVARLSQRPKAEFAEYVKLVITDLTK 240  
 Db 392 KLDELREGKASSAKORLKCSLQKFGERAFKAWAVARLSQRPKAEFAEYVKLVITDLTK 451  
 QY 241 VHTCCCHGDLLECADRADLAKYICENODSISSKLKECCCKPFLLEKSHCIAEVENDEMPA 300  
 Db 452 VHTCCCHGDLLECADRADLAKYICENODSISSKLKECCCKPFLLEKSHCIAEVENDEMPA 511  
 QY 301 DLPSLAADFVESKDVCKNTAEAKOVFLGMFLYETARRHPDYSVVLLRLAKTYETTTLEK 360  
 Db 512 DLPSLAADFVESKDVCKNTAEAKOVFLGMFLYETARRHPDYSVVLLRLAKTYETTTLEK 571  
 QY 361 CAADPEHCYAKVDFEDEFKPLVEEPONLIKONCELFQOLGEYFQNALVRYTKKVPQVST 420  
 Db 572 CAADPEHCYAKVDFEDEFKPLVEEPONLIKONCELFQOLGEYFQNALVRYTKKVPQVST 631  
 QY 421 PTLVEVSRLNGKVGSKCKKHPKAMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
 Db 632 PTLVEVSRLNGKVGSKCKKHPKAMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 691  
 QY 481 LVNRRPCFSALEVDETYVPEKFNATFTFHADICTLSEKERQIKKQTALVELVHKHPKAT 540  
 Db 692 LVNRRPCFSALEVDETYVPEKFNATFTFHADICTLSEKERQIKKQTALVELVHKHPKAT 751  
 QY 541 KEOLKAYMDDFAAFVEKCKCKADDDKTCFAEBGKKLVAASQAALG 584  
 Db 752 KEOLKAYMDDFAAFVEKCKCKADDDKTCFAEBGKKLVAASQAALG 795

## RESULT 13

US-08-448-196A-3  
 ; Sequence 3, Application US/08448196A  
 ; Patent No. 5780594  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CARTER, DANIEL C.  
 ; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
 ; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
 ; TITLE OF INVENTION: RELATED PROTEINS  
 ; NUMBER OF SEQUENCES: 9  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: NASA  
 ; STREET: MARSHALL SPACE FLIGHT CENTER  
 ; CITY: HUNTSVILLE  
 ; STATE: ALABAMA  
 ; COUNTRY: USA  
 ; ZIP: 35812  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/448,196A  
 ; FILING DATE: 23-MAY-1995  
 ; CLASSIFICATION: 530



## ATTORNEY/AGENT INFORMATION:

NAME: BROAD JR., ROBERT L.  
 REGISTRATION NUMBER: 18,757  
 REFERENCE/DOCKET NUMBER: XX/MFS-28402-2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 205-544-0021  
 TELEFAX: 205-544-0258

## INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 585 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHEICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: N-terminal

US-08-448-196A-3

Query Match 99.7%; Score 3093; DB 1; Length 585;  
 Best Local Similarity 99.7%; Pred. No. 8.3e-286;  
 Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLOQCPFDHVKLVNEVTEFAKTCVADESAE 60  
 DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLOQCPFDHVKLVNEVTEFAKTCVADESAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVPRPV 120  
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVPRPV 120

QY 121 DVMTAFHDNEETFLKLYIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180  
 DB 121 DVMTAFHDNEETFLKLYIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180

QY 181 KLDELDEGKASSAKQRLKCSAQKGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240  
 DB 181 KLDELDEGKASSAKQRLKCSAQKGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240

QY 241 VHTCCGDLLEACADRADLAKYICENODSISSKLKECCERPLEKSHCIAEVNDEMFA 300  
 DB 241 VHTCCGDLLEACADRADLAKYICENODSISSKLKECCERPLEKSHCIAEVNDEMFA 300

QY 301 DLPFLAADVESEKDYCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLRLAKTYETTLK 360  
 DB 301 DLPFLAADVESEKDYCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLRLAKTYETTLK 360

QY 361 CAAADPHECYAKVDFEFPPLVEEPQNLKQCELPQGEYKFNALLVYTKKVPQVST 420  
 DB 361 CAAADPHECYAKVDFEFPPLVEEPQNLKQCELPQGEYKFNALLVYTKKVPQVST 420

QY 421 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480  
 DB 421 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480

QY 481 LVNRRPCFSALEVDYETYPKFNATFTFHADICTLSEKEROIKQKOTALVELVKKPKAT 540  
 DB 481 LVNRRPCFSALEVDYETYPKFNATFTFHADICTLSEKEROIKQKOTALVELVKKPKAT 540

QY 541 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKKLVAAASQAALGL 585  
 DB 541 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKKLVAAASQAALGL 585

## RESULT 14

US-08-984-176-1  
 Sequence 1, Application US/08984176  
 Patent No. 5948609

## GENERAL INFORMATION:

APPLICANT: DANIEL C  
 APPLICANT: HO, JOSEPH X  
 APPLICANT: KUKER, FLORIAN  
 TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT  
 TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER

FILE REFERENCE: 08/984,176  
 CURRENT APPLICATION NUMBER: US/08/984,176  
 CURRENT FILING DATE: 1997-12-03

NUMBER OF SEQ ID NOS: 1

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1

LENGTH: 585

TYPE: PRT

ORGANISM: Homo sapiens

US-08-984-176-1

Query Match 99.7%; Score 3093; DB 2; Length 585;  
 Best Local Similarity 99.7%; Pred. No. 8.3e-286;  
 Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLOQCPFDHVKLVNEVTEFAKTCVADESAE 60  
 DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLOQCPFDHVKLVNEVTEFAKTCVADESAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVPRPV 120  
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVPRPV 120

QY 121 DVMTAFHDNEETFLKLYIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180  
 DB 121 DVMTAFHDNEETFLKLYIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180

QY 181 KLDELDEGKASSAKQRLKCSAQKGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240  
 DB 181 KLDELDEGKASSAKQRLKCSAQKGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240

QY 241 VHTCCGDLLEACADRADLAKYICENODSISSKLKECCERPLEKSHCIAEVNDEMFA 300  
 DB 241 VHTCCGDLLEACADRADLAKYICENODSISSKLKECCERPLEKSHCIAEVNDEMFA 300

QY 301 DLPFLAADVESEKDYCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLRLAKTYETTLK 360  
 DB 301 DLPFLAADVESEKDYCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLRLAKTYETTLK 360

QY 361 CAAADPHECYAKVDFEFPPLVEEPQNLKQCELPQGEYKFNALLVYTKKVPQVST 420  
 DB 361 CAAADPHECYAKVDFEFPPLVEEPQNLKQCELPQGEYKFNALLVYTKKVPQVST 420

QY 421 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480  
 DB 421 PTLVEYSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480

QY 481 LVNRRPCFSALEVDYETYPKFNATFTFHADICTLSEKEROIKQKOTALVELVKKPKAT 540  
 DB 481 LVNRRPCFSALEVDYETYPKFNATFTFHADICTLSEKEROIKQKOTALVELVKKPKAT 540

QY 541 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKKLVAAASQAALGL 585  
 DB 541 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKKLVAAASQAALGL 585

## RESULT 15

US-08-448-196A-5

Sequence 5, Application US/08448196A

Patent No. 5780594

GENERAL INFORMATION:

APPLICANT: CARTER, DANIEL C.

TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS

TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR

TITLE OF INVENTION: RELATED PROTEINS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: NASA

STREET: MARSHALL SPACE FLIGHT CENTER

CITY: HUNTSVILLE

STATE: ALABAMA

COUNTRY: USA

ZIP: 35812

Job time : 32 secs

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/448,196A  
FILING DATE: 23-MAY-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: BROAD JR., ROBERT L.  
REGISTRATION NUMBER: 18,757  
REFERENCE/DOCKET NUMBER: XX/MFS-28402-2  
TELEPHONE: 205-544-0021  
TELEFAX: 205-544-0258  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 583 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-08-448-196A-5

Query Match 79.2%; Score 2458.5; DB 1; Length 583;  
Best local Similarity 75.8%; Pred. No. 2.1e-225;  
Matches 442; Conservative 70; Mismatches 70; Indels 1; Gaps 1;

QY	1	DAHKEVAREKDLGEENFALVLIARFAYLIQOCPPEDHVKLVNEVTEFAKTCVADSEAE	60
Db	1	DTHKSEIARRNDLGEARFGLVAFSQTQQPPEDHVKLVNEVTEFAKTCVADSEAE	60
QY	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPLRVPEV	120
Db	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPLRVPEV	119
QY	121	DVMTAFHDNBEFLKYLVEIARRHPYAPPELLFAKRYKAAFTCCQAADKAACLLP	180
Db	120	DAQAAFDQDPKFLGLYEVARRHPYFYPPELLFAEYKADFTCCPADDKIACLIP	179
QY	181	KLDELDFDEGKASSAKORLKASLOKTERAFKANAVALRSORPPKAEFAEVSCLVDTLK	240
Db	180	KLDALAKERTILLSAKERLKCSSPQNGEGRANVANSVALSQPKADPAEVSKIYDTLK	239
QY	241	VHPECCHGLLECADRADLAKYICENQDSISSKKECKPILLEKSHCHIAEVNDEMPA	300
Db	240	VHPECCHGLLECADRADLAKYICENQDSISSKKECKPILLEKSHCHIAEVNDEMPA	299
QY	301	DLPSLAADPVESKOVCKNYAEKDVFLGMFLYEVARRHPDYSVLLIRLAKTYETTLKC	360
Db	300	DIPALAADPAEDKEICRYKADKDVFLGFLYEVARRHPDYSVLLIRLAKTYETTLKC	359
QY	361	CAAADPHECYAKVDEFEKPLVEBPQNLKQNCLEPQGEYKFNALLVRYTKKVPQVST	420
Db	360	CAEADPPACRYTVEDQFTPLVEBPKSLVKNCDLFEVEGYEDFONALIVRYTKKAPQVST	419
QY	421	PTLVFVSRLGKVGSKCKHPKAPCAEDYLSVNLQICVLHEKTPVSDRYTKCTES	480
Db	420	PTLVETGRILGKVGSRCKLPESERLPCSENHALALNRLCVLHEKTPVSEKITCTCDS	479
QY	481	LVNRRPFCFSALEVDVTVYPKEFNAETFTFHADICTLSEKROIKKOTALVELVKKPKAT	540
Db	480	LAERRCFCFSALEDDGTPVKEFAETFTFHADICTLPEDEKQIKKQSAALAEVKKPKAT	539
QY	541	KEQLKAVMDDFAAFEYKCKADDKETCFABEGCKLVAASQAAL	583
Db	540	KEQLKTVLGNFSAFVAKCGREDKEACFAEAGGPKLVASSQAL	582

Search completed: August 31, 2003, 16:37:08

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2003, 16:33:53 ; Search time 28 Seconds  
(without alignments)  
2857.764 Million cell updates/sec

Title: US-09-833-041-18

Perfect score: 3103  
Sequence: 1 DAKSEVAHRFKDGLGFNFK.....TCFAEGKKLVAASQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 510680 seqs, 136781880 residues  
Total number of hits satisfying chosen parameters: 510680

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lasting first 45 summaries

Database : Published Applications AA:\*

1: /cgn2.6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2.6/ptodata/2/pubpaa/PTCT\_NEW\_PUB.pep.\*  
3: /cgn2.6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2.6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2.6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2.6/ptodata/2/pubpaa/PTCTUS\_PUBCOMB.pep.\*  
7: /cgn2.6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2.6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2.6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2.6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2.6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2.6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2.6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2.6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2.6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2.6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2.6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2.6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	3103	100.0	585	10	US-09-929-552-2 Sequence 2, Appli
2	3103	100.0	585	11	US-09-932-613-445 Sequence 445, App
3	3103	100.0	585	11	US-09-984-010-26 Sequence 26, Appl
4	3103	100.0	585	11	US-09-833-041-18 Sequence 18, Appl
5	3103	100.0	585	12	US-10-153-604A-5 Sequence 5, Appli
6	3103	100.0	585	14	US-10-153-064-5 Sequence 5, Appli
7	3103	100.0	604	11	US-09-984-010-7 Sequence 7, Appli
8	3103	100.0	609	11	US-09-919-039-370 Sequence 370, App
9	3103	100.0	609	12	US-10-153-604A-7 Sequence 7, Appli
10	3103	100.0	609	14	US-10-153-064-7 Sequence 7, Appli
11	3103	100.0	610	10	US-09-984-186-2 Sequence 2, Appli
12	3103	100.0	610	15	US-10-237-667-2 Sequence 2, Appli
13	3103	100.0	610	15	US-10-237-708-2 Sequence 2, Appli
14	3103	100.0	610	15	US-10-237-866-2 Sequence 2, Appli
15	3103	100.0	610	15	US-10-237-871-2 Sequence 2, Appli

Sequence 2, Appli  
Sequence 133, App  
Sequence 133, App  
Sequence 132, App  
Sequence 132, App  
Sequence 132, App  
Sequence 131, App  
Sequence 131, App  
Sequence 130, App  
Sequence 130, App  
Sequence 127, App  
Sequence 127, App  
Sequence 127, App  
Sequence 127, App  
Sequence 125, App  
Sequence 125, App  
Sequence 123, App  
Sequence 123, App  
Sequence 123, App  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 96, Appl  
Sequence 96, Appl  
Sequence 99, Appl  
Sequence 105, App  
Sequence 96, App  
Sequence 99, App  
Sequence 105, App

16 3103 100.0 610 15 US-10-237-624-2  
17 3103 100.0 651 12 US-10-153-604A-133  
18 3103 100.0 651 14 US-10-153-064-133  
19 3103 100.0 652 12 US-10-153-604A-132  
20 3103 100.0 652 14 US-10-153-064-132  
21 3103 100.0 653 12 US-10-153-604A-131  
22 3103 100.0 653 14 US-10-153-064-131  
23 3103 100.0 656 12 US-10-153-604A-130  
24 3103 100.0 656 14 US-10-153-064-130  
25 3103 100.0 676 12 US-10-153-604A-127  
26 3103 100.0 676 12 US-10-153-604A-129  
27 3103 100.0 676 14 US-10-153-064-127  
28 3103 100.0 676 12 US-10-153-064-129  
29 3103 100.0 677 14 US-10-153-604A-125  
30 3103 100.0 677 12 US-10-153-064-125  
31 3103 100.0 680 12 US-10-153-604A-123  
32 3103 100.0 680 14 US-10-153-064-123  
33 3103 100.0 787 10 US-09-984-186-16  
34 3103 100.0 787 15 US-10-237-667-16  
35 3103 100.0 787 15 US-10-237-708-16  
36 3103 100.0 787 15 US-10-237-866-16  
37 3103 100.0 787 15 US-10-237-871-16  
38 3103 100.0 787 15 US-10-237-624-16  
39 3103 100.0 788 15 US-10-073-118-26  
40 3092.5 99.7 652 12 US-10-153-604A-96  
41 3092.5 99.7 652 12 US-10-153-604A-98  
42 3092.5 99.7 652 12 US-10-153-604A-105  
43 3092.5 99.7 652 14 US-10-153-064-96  
44 3092.5 99.7 652 14 US-10-153-064-99  
45 3092.5 99.7 652 14 US-10-153-064-105

#### ALIGNMENTS

RESULT 1  
US-09-929-552-2  
Sequence 2, Application US/0929552  
Patent No. US20020123080A1  
GENERAL INFORMATION:  
APPLICANT: Sonnenschein, Carlos  
Soto, Ana M.  
TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Medlen & Carroll, LLP  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/929,552  
FILING DATE: 14-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/769,746  
FILING DATE: 19-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Peter G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: MBRI-02584  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 585 amino acids

; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-09-929-552-2

Query Match 100.0%; Score 3103; DB 10; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 1e-270;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DAHSEVAHRFDLGEENFKALVLIAPQAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
 Db 1 DAHSEVAHRFDLGEENFKALVLIAPQAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
 QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVPEV 120  
 Db 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVPEV 120  
 QY 121 DVMCTAPHDNEETFLKKLYIEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180  
 Db 121 DVMCTAPHDNEETFLKKLYIEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180  
 QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVPEV 120  
 Db 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVPEV 120  
 QY 121 DVMCTAPHDNEETFLKKLYIEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180  
 Db 121 DVMCTAPHDNEETFLKKLYIEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180  
 QY 181 KLDELDEGKASSAKORLKCSLOKFGERAFKAWARLSQRFPAEFAEVS KLVTDLTK 240  
 Db 181 KLDELDEGKASSAKORLKCSLOKFGERAFKAWARLSQRFPAEFAEVS KLVTDLTK 240  
 QY 181 KLDELDEGKASSAKORLKCSLOKFGERAFKAWARLSQRFPAEFAEVS KLVTDLTK 240  
 Db 181 KLDELDEGKASSAKORLKCSLOKFGERAFKAWARLSQRFPAEFAEVS KLVTDLTK 240  
 QY 241 VTECCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
 Db 241 VTECCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
 QY 301 DLPSLAADRVESKDVCKNYAEAKOVFLGMFLYIYARRHPDYSVVLLRLAKTYETTLK 360  
 Db 301 DLPSLAADRVESKDVCKNYAEAKOVFLGMFLYIYARRHPDYSVVLLRLAKTYETTLK 360  
 QY 361 CAADAPHECVAKVDFEFLPEEONLIKONCELFOLGEYKFNALLVRYTKKVPQVST 420  
 Db 361 CAADAPHECVAKVDFEFLPEEONLIKONCELFOLGEYKFNALLVRYTKKVPQVST 420  
 QY 301 DLPSLAADRVESKDVCKNYAEAKOVFLGMFLYIYARRHPDYSVVLLRLAKTYETTLK 360  
 Db 301 DLPSLAADRVESKDVCKNYAEAKOVFLGMFLYIYARRHPDYSVVLLRLAKTYETTLK 360  
 QY 361 CAADAPHECVAKVDFEFLPEEONLIKONCELFOLGEYKFNALLVRYTKKVPQVST 420  
 Db 361 CAADAPHECVAKVDFEFLPEEONLIKONCELFOLGEYKFNALLVRYTKKVPQVST 420  
 QY 421 PTLVEVSRNLGVKSGCKCKHPKAMPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCOTES 480  
 Db 421 PTLVEVSRNLGVKSGCKCKHPKAMPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCOTES 480  
 QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKQTALVELVYKHFKPAT 540  
 Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKQTALVELVYKHFKPAT 540  
 QY 541 KEOLKAVMDFFAFAVEKCKCKADDDKTCFAEAGKLVAAASQAALGL 585  
 Db 541 KEOLKAVMDFFAFAVEKCKCKADDDKTCFAEAGKLVAAASQAALGL 585

RESULT 2  
 US-09-932-613-445  
 ; Sequence 445, Application US/09932613  
 ; Publication No. US20030091565A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Human Genome Sciences, Inc.  
 ; APPLICANT: Belter, James P.  
 ; APPLICANT: Potter, M. Daniel  
 ; APPLICANT: Fleming, Tony J.  
 ; APPLICANT: Rosen, Craig A.  
 ; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON  
 ; FILE REFERENCE: DXX-025.1 PCT; DXX-025.1 US  
 ; CURRENT FILING DATE: 2001-08-17  
 ; NUMBER OF SEQ ID NOS: 458  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 445  
 ; LENGTH: 585  
 ; TYPE: PRT  
 ; ORGANISM: HomoSapiens  
 US-09-932-613-445

Query Match 100.0%; Score 3103; DB 11; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 1e-270;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DAHSEVAHRFDLGEENFKALVLIAPQAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
 Db 1 DAHSEVAHRFDLGEENFKALVLIAPQAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60  
 QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVPEV 120  
 Db 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVPEV 120  
 QY 121 DVMCTAPHDNEETFLKKLYIEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180  
 Db 121 DVMCTAPHDNEETFLKKLYIEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180  
 QY 181 KLDELDEGKASSAKORLKCSLOKFGERAFKAWARLSQRFPAEFAEVS KLVTDLTK 240  
 Db 181 KLDELDEGKASSAKORLKCSLOKFGERAFKAWARLSQRFPAEFAEVS KLVTDLTK 240  
 QY 241 VTECCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
 Db 241 VTECCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
 QY 301 DLPSLAADRVESKDVCKNYAEAKOVFLGMFLYIYARRHPDYSVVLLRLAKTYETTLK 360  
 Db 301 DLPSLAADRVESKDVCKNYAEAKOVFLGMFLYIYARRHPDYSVVLLRLAKTYETTLK 360  
 QY 361 CAADAPHECVAKVDFEFLPEEONLIKONCELFOLGEYKFNALLVRYTKKVPQVST 420  
 Db 361 CAADAPHECVAKVDFEFLPEEONLIKONCELFOLGEYKFNALLVRYTKKVPQVST 420  
 QY 421 PTLVEVSRNLGVKSGCKCKHPKAMPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCOTES 480  
 Db 421 PTLVEVSRNLGVKSGCKCKHPKAMPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCOTES 480  
 QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKQTALVELVYKHFKPAT 540  
 Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKQTALVELVYKHFKPAT 540  
 QY 541 KEOLKAVMDFFAFAVEKCKCKADDDKTCFAEAGKLVAAASQAALGL 585  
 Db 541 KEOLKAVMDFFAFAVEKCKCKADDDKTCFAEAGKLVAAASQAALGL 585

RESULT 3  
 US-09-984-010-26  
 ; Sequence 26, Application US/09984010  
 ; Publication No. US20030104578A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ballance, David James  
 ; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE AND SERUM ALBUMIN  
 ; NUMBER OF SEQUENCES: 26  
 ; CORRESPONDENCE ADDRESSES:  
 ; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP  
 ; STREET: 1300 I Street, NW  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20005-3315  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/984.010  
 ; FILING DATE: 21-May-2002  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 09/091,873  
 ; FILING DATE: 25-JUN-1998  
 ; APPLICATION NUMBER: PCT/GB96/03164

/ FILING DATE: 19-DEC-1996  
/ INFORMATION FOR SEQ ID NO: 26:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 585 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS: <unknown>  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: protein  
/ HYPOTHEetical: NO  
/ ANTI-SENSE: NO  
/ SEQUENCE DESCRIPTION: SEQ ID NO: 26:  
US-09-984-010-26

Query Match 100.0%; Score 3103; DB 11; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1e-270;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIAPFAQYLOCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHFRKDLGEENFKALVLIAPFAQYLOCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVREPV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVREPV 120

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVREPV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVREPV 120

QY 121 DVMTAFHNEETFLKKLYLFIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
DB 121 DVMTAFHNEETFLKKLYLFIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

QY 121 DVMTAFHNEETFLKKLYLFIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
DB 121 DVMTAFHNEETFLKKLYLFIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRFPAEFAEVSRLVTDLT 240  
DB 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRFPAEFAEVSRLVTDLT 240

QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRFPAEFAEVSRLVTDLT 240  
DB 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRFPAEFAEVSRLVTDLT 240

QY 241 VHTCCGHDLLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCCGHDLLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300

QY 241 VHTCCGHDLLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCCGHDLLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300

QY 301 DLPFLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360  
DB 301 DLPFLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360

QY 301 DLPFLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360  
DB 301 DLPFLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360

QY 361 CAADPHECYAKVDFEFKPLVEEPQNLKQNCSELFQOLGEYKFNALLVRYTKKPOVST 420  
DB 361 CAADPHECYAKVDFEFKPLVEEPQNLKQNCSELFQOLGEYKFNALLVRYTKKPOVST 420

QY 361 CAADPHECYAKVDFEFKPLVEEPQNLKQNCSELFQOLGEYKFNALLVRYTKKPOVST 420  
DB 361 CAADPHECYAKVDFEFKPLVEEPQNLKQNCSELFQOLGEYKFNALLVRYTKKPOVST 420

QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480

QY 481 LVNRRPCFSALEVDYTYVPKEFNAETFTFHADICTLSEKEROIKQTALVELVKKPKAT 540  
DB 481 LVNRRPCFSALEVDYTYVPKEFNAETFTFHADICTLSEKEROIKQTALVELVKKPKAT 540

QY 541 KEQLKAVMDFFAAFEVKCKCKADDDKTCFAEKGKLVAAASQAALGL 585  
DB 541 KEQLKAVMDFFAAFEVKCKCKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 4  
US-09-833-041-18

/ Sequence 18, Application US/09833041  
/ Publication NO. US20030125247A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Rosen, Craig A.  
/ APPLICANT: Haseltine, William A.  
/ TITLE OF INVENTION: Albumin Fusion Proteins  
/ FILE REFERENCE: PF545  
/ CURRENT APPLICATION NUMBER: US/09/833,041  
/ PRIOR FILING DATE: 2001-04-12  
/ SEQUENCE DESCRIPTION: SEQ ID NO: 26:  
US-09-984-010-26

/ PRIOR APPLICATION NUMBER: 60/199,384  
/ PRIOR FILING DATE: 2000-04-25  
/ NUMBER OF SEQ ID NOS: 79  
/ SOFTWARE: PatentIn Ver. 2.1  
/ SEQ ID NO 18  
/ LENGTH: 585  
/ TYPE: PRT  
/ ORGANISM: Homo Sapiens  
US-09-833-041-18

Query Match 100.0%; Score 3103; DB 11; Length 585;  
Best Local Similarity 100.0%; Pred. No. 1e-270;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIAPFAQYLOCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHFRKDLGEENFKALVLIAPFAQYLOCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVREPV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVREPV 120

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVREPV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVREPV 120

QY 121 DVMTAFHNEETFLKKLYLFIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
DB 121 DVMTAFHNEETFLKKLYLFIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

QY 121 DVMTAFHNEETFLKKLYLFIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180  
DB 121 DVMTAFHNEETFLKKLYLFIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRFPAEFAEVSRLVTDLT 240  
DB 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRFPAEFAEVSRLVTDLT 240

QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRFPAEFAEVSRLVTDLT 240  
DB 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRFPAEFAEVSRLVTDLT 240

QY 241 VHTCCGHDLLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCCGHDLLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300

QY 241 VHTCCGHDLLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCCGHDLLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300

QY 301 DLPFLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360  
DB 301 DLPFLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360

QY 301 DLPFLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360  
DB 301 DLPFLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLRLAKTYETTLK 360

QY 361 CAADPHECYAKVDFEFKPLVEEPQNLKQNCSELFQOLGEYKFNALLVRYTKKPOVST 420  
DB 361 CAADPHECYAKVDFEFKPLVEEPQNLKQNCSELFQOLGEYKFNALLVRYTKKPOVST 420

QY 361 CAADPHECYAKVDFEFKPLVEEPQNLKQNCSELFQOLGEYKFNALLVRYTKKPOVST 420  
DB 361 CAADPHECYAKVDFEFKPLVEEPQNLKQNCSELFQOLGEYKFNALLVRYTKKPOVST 420

QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480  
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480

QY 481 LVNRRPCFSALEVDYTYVPKEFNAETFTFHADICTLSEKEROIKQTALVELVKKPKAT 540  
DB 481 LVNRRPCFSALEVDYTYVPKEFNAETFTFHADICTLSEKEROIKQTALVELVKKPKAT 540

QY 541 KEQLKAVMDFFAAFEVKCKCKADDDKTCFAEKGKLVAAASQAALGL 585  
DB 541 KEQLKAVMDFFAAFEVKCKCKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 5

US-10-153-604A-5  
/ Sequence 5, Application US/10153604A  
/ Publication NO. US20030143191A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Bell et al.  
/ TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
/ FILE REFERENCE: PF556  
/ CURRENT APPLICATION NUMBER: US/10/153,604A  
/ PRIOR FILING DATE: 2002-05-24  
/ SEQUENCE DESCRIPTION: SEQ ID NO: 137  
US-10-153-604A-5

US-10-153-604A-5

Query Match 100.0%; Score 3103; DB 12; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 1e-270;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREFKDLGSENFKALVLIATAFYAQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60  
 |||||  
 Db 1 DAHSEVAHREFKDLGSENFKALVLIATAFYAQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60  
 |||||  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKPEPERNECFLOHKDDNPMLPLRVPEV 120  
 |||||  
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKPEPERNECFLOHKDDNPMLPLRVPEV 120  
 |||||  
 QY 121 DVMCTAFDNEETFLKKYLYEIARHPYFYAPPELLFFAKRYKAAFTCCQAADKAACILP 180  
 |||||  
 Db 121 DVMCTAFDNEETFLKKYLYEIARHPYFYAPPELLFFAKRYKAAFTCCQAADKAACILP 180  
 |||||  
 QY 181 KIDELRDEGKASSAKORLKCASLOKFGGERAFKAWARLSQRPFAEFAEYVSKLVTDLTK 240  
 |||||  
 Db 181 KIDELRDEGKASSAKORLKCASLOKFGGERAFKAWARLSQRPFAEFAEYVSKLVTDLTK 240  
 |||||  
 QY 241 VHTCCCHGDLLECADRDLAKYICENODSISSKLKCECKPLEKSKHSCIAEVENDEMPA 300  
 |||||  
 Db 241 VHTCCCHGDLLECADRDLAKYICENODSISSKLKCECKPLEKSKHSCIAEVENDEMPA 300  
 |||||  
 QY 301 DLPSLAADFVESKDVCKNTAEAKDVFLGMFLYFYARHPDYVSVLLRLAKTYETTLTK 360  
 |||||  
 Db 301 DLPSLAADFVESKDVCKNTAEAKDVFLGMFLYFYARHPDYVSVLLRLAKTYETTLTK 360  
 |||||  
 QY 361 CAADPHECYAKVDFEFKPLVEEPONLTKONCELFEOLOGYFQNALVRLVTKKVPQVST 420  
 |||||  
 Db 361 CAADPHECYAKVDFEFKPLVEEPONLTKONCELFEOLOGYFQNALVRLVTKKVPQVST 420  
 |||||  
 QY 421 PTLVEVSRNLGKVGSKCKKHPKAPCAEDYLSVNLNOLCVLHKTVPVSDRVTKCCTES 480  
 |||||  
 Db 421 PTLVEVSRNLGKVGSKCKKHPKAPCAEDYLSVNLNOLCVLHKTVPVSDRVTKCCTES 480  
 |||||  
 QY 481 LVNRRPCFSALEVDYTYVPKEFNAETFFHADICTLSEKEROIKKOTALVELVHKPKAT 540  
 |||||  
 Db 481 LVNRRPCFSALEVDYTYVPKEFNAETFFHADICTLSEKEROIKKOTALVELVHKPKAT 540  
 |||||  
 QY 541 KEOLKAVMDFAAFVEKCKKADKTCFAEGBGKKLVAASQAALGL 585  
 |||||  
 Db 541 KEOLKAVMDFAAFVEKCKKADKTCFAEGBGKKLVAASQAALGL 585  
 |||||

## RESULT 6

US-10-153-064-5  
 ; Sequence 5, Application US/10153064  
 ; Publication No. US20020142814A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bell et al.  
 ; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
 ; FILE REFERENCE: PF556  
 ; CURRENT APPLICATION NUMBER: US/10/153,064  
 ; PRIOR FILING DATE: 2002-05-24  
 ; PRIOR APPLICATION NUMBER: 60/293,212  
 ; PRIOR FILING DATE: 2001-05-25  
 ; NUMBER OF SEQ ID NOS: 137  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 5  
 ; LENGTH: 585  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapiens  
 US-10-153-064-5

Query Match 100.0%; Score 3103; DB 14; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 1e-270;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREFKDLGSENFKALVLIATAFYAQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60  
 |||||  
 |||||

Db 1 DAHSEVAHREFKDLGSENFKALVLIATAFYAQYLQCCPFEDHVKLVNEVTEFAKTCVADSSAE 60  
 |||||  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKPEPERNECFLOHKDDNPMLPLRVPEV 120  
 |||||  
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKPEPERNECFLOHKDDNPMLPLRVPEV 120  
 |||||  
 QY 121 DVMCTAFDNEETFLKKYLYEIARHPYFYAPPELLFFAKRYKAAFTCCQAADKAACILP 180  
 |||||  
 Db 121 DVMCTAFDNEETFLKKYLYEIARHPYFYAPPELLFFAKRYKAAFTCCQAADKAACILP 180  
 |||||  
 QY 181 KIDELRDEGKASSAKORLKCASLOKFGGERAFKAWARLSQRPFAEFAEYVSKLVTDLTK 240  
 |||||  
 Db 181 KIDELRDEGKASSAKORLKCASLOKFGGERAFKAWARLSQRPFAEFAEYVSKLVTDLTK 240  
 |||||  
 QY 241 VHTCCCHGDLLECADRDLAKYICENODSISSKLKCECKPLEKSKHSCIAEVENDEMPA 300  
 |||||  
 Db 241 VHTCCCHGDLLECADRDLAKYICENODSISSKLKCECKPLEKSKHSCIAEVENDEMPA 300  
 |||||  
 QY 301 DLPSLAADFVESKDVCKNTAEAKDVFLGMFLYFYARHPDYVSVLLRLAKTYETTLTK 360  
 |||||  
 Db 301 DLPSLAADFVESKDVCKNTAEAKDVFLGMFLYFYARHPDYVSVLLRLAKTYETTLTK 360  
 |||||  
 QY 361 CAADPHECYAKVDFEFKPLVEEPONLTKONCELFEOLOGYFQNALVRLVTKKVPQVST 420  
 |||||  
 Db 361 CAADPHECYAKVDFEFKPLVEEPONLTKONCELFEOLOGYFQNALVRLVTKKVPQVST 420  
 |||||  
 QY 421 PTLVEVSRNLGKVGSKCKKHPKAPCAEDYLSVNLNOLCVLHKTVPVSDRVTKCCTES 480  
 |||||  
 Db 421 PTLVEVSRNLGKVGSKCKKHPKAPCAEDYLSVNLNOLCVLHKTVPVSDRVTKCCTES 480  
 |||||  
 QY 481 LVNRRPCFSALEVDYTYVPKEFNAETFFHADICTLSEKEROIKKOTALVELVHKPKAT 540  
 |||||  
 Db 481 LVNRRPCFSALEVDYTYVPKEFNAETFFHADICTLSEKEROIKKOTALVELVHKPKAT 540  
 |||||  
 QY 541 KEOLKAVMDFAAFVEKCKKADKTCFAEGBGKKLVAASQAALGL 585  
 |||||  
 Db 541 KEOLKAVMDFAAFVEKCKKADKTCFAEGBGKKLVAASQAALGL 585  
 |||||

## RESULT 7

US-09-984-010-7  
 ; Sequence 7, Application US/09984010  
 ; Publication No. US20030104578A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ballance, David James  
 ; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE  
 ; AND SERUM ALBUMIN  
 ; NUMBER OF SEQUENCES: 26  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP  
 ; STREET: 1300 I Street, NW  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20005-3315  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
 ; CURRENT APPLICATION NUMBER: US/09/984,010  
 ; FILING DATE: 21-May-2002  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 09/091,873  
 ; FILING DATE: 25-JUN-1998  
 ; APPLICATION NUMBER: PCT/GB96/03164  
 ; FILING DATE: 19-DEC-1996  
 ; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 604 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: <Unknown>

```
;
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-984-010-7

Query Match      100.0%; Score 3103; DB 11; Length 604;
Best Local Similarity 100.0%; Pred. No. 1.le-270;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFDLGEENFKALVLIAPAFYLOQCFFEDHVKLVNVEYFAKTCVADSAE 60
   |||||||
Db 20 DAHKSVAHRFDLGEENFKALVLIAPAFYLOQCFFEDHVKLVNVEYFAKTCVADSAE 79
   |||||||

QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPNLRLVRPEV 120
   |||||||
Db 80 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPNLRLVRPEV 139
   |||||||

QY 121 DVMTAFHDNEETFLKKYLXIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
   |||||||
Db 140 DVMTAFHDNEETFLKKYLXIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 199
   |||||||

QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSORFPKAEFVSKLVTDLTK 240
   |||||||
Db 200 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSORFPKAEFVSKLVTDLTK 259
   |||||||

QY 241 VHECCGDLLECCADRADLAKYICENQDSISSKLKECCERPLEKSHCIAEVENDEMPA 300
   |||||||
Db 260 VHECCGDLLECCADRADLAKYICENQDSISSKLKECCERPLEKSHCIAEVENDEMPA 319
   |||||||

QY 301 DLPSLAADFVESKDVCKNYAKDVFLGMFLYEYARRHPDYSVLLLRKATYTTLEKC 360
   |||||||
Db 320 DLPSLAADFVESKDVCKNYAKDVFLGMFLYEYARRHPDYSVLLLRKATYTTLEKC 379
   |||||||

QY 361 CAADAPHECYAKVDFEFLVEEPQNLIKONCELPOLGEYKFNALLVRYTKKVPQVST 420
   |||||||
Db 380 CAADAPHECYAKVDFEFLVEEPQNLIKONCELPOLGEYKFNALLVRYTKKVPQVST 439
   |||||||

QY 421 PTLVEYSRNLKGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
   |||||||
Db 440 PTLVEYSRNLKGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 499
   |||||||

QY 481 LVNRPCFSALVDETYTPKFEFNAETFTFHADICTLSEKERQIKKQALVELYKHKPKAT 540
   |||||||
Db 500 LVNRPCFSALVDETYTPKFEFNAETFTFHADICTLSEKERQIKKQALVELYKHKPKAT 559
   |||||||

QY 541 KEQLKAVMDDFAAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL 585
   |||||||
Db 560 KEQLKAVMDDFAAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL 604
   |||||||
```

```
RESULT 8
US-09-919-039-370
; Sequence 370, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; PRIOR FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 370
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 088957CD1
```

```
US-09-919-039-370

Query Match      100.0%; Score 3103; DB 11; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.le-270;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFDLGEENFKALVLIAPAFYLOQCFFEDHVKLVNVEYFAKTCVADSAE 60
   |||||||
Db 25 DAHKSVAHRFDLGEENFKALVLIAPAFYLOQCFFEDHVKLVNVEYFAKTCVADSAE 84
   |||||||

QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPNLRLVRPEV 120
   |||||||
Db 85 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPNLRLVRPEV 144
   |||||||

QY 121 DVMTAFHDNEETFLKKYLXIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
   |||||||
Db 145 DVMTAFHDNEETFLKKYLXIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 204
   |||||||

QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSORFPKAEFVSKLVTDLTK 240
   |||||||
Db 205 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSORFPKAEFVSKLVTDLTK 264
   |||||||

QY 241 VHECCGDLLECCADRADLAKYICENQDSISSKLKECCERPLEKSHCIAEVENDEMPA 300
   |||||||
Db 265 VHECCGDLLECCADRADLAKYICENQDSISSKLKECCERPLEKSHCIAEVENDEMPA 324
   |||||||

QY 301 DLPSLAADFVESKDVCKNYAKDVFLGMFLYEYARRHPDYSVLLLRKATYTTLEKC 360
   |||||||
Db 325 DLPSLAADFVESKDVCKNYAKDVFLGMFLYEYARRHPDYSVLLLRKATYTTLEKC 384
   |||||||

QY 361 CAADAPHECYAKVDFEFLVEEPQNLIKONCELPOLGEYKFNALLVRYTKKVPQVST 420
   |||||||
Db 385 CAADAPHECYAKVDFEFLVEEPQNLIKONCELPOLGEYKFNALLVRYTKKVPQVST 444
   |||||||

QY 421 PTLVEYSRNLKGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
   |||||||
Db 445 PTLVEYSRNLKGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 504
   |||||||

QY 481 LVNRPCFSALVDETYTPKFEFNAETFTFHADICTLSEKERQIKKQALVELYKHKPKAT 540
   |||||||
Db 505 LVNRPCFSALVDETYTPKFEFNAETFTFHADICTLSEKERQIKKQALVELYKHKPKAT 564
   |||||||

QY 541 KEQLKAVMDDFAAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL 585
   |||||||
Db 565 KEQLKAVMDDFAAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL 609
   |||||||
```

```
RESULT 9
US-10-153-604A-7
; Sequence 7, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,604A
; CURRENT FILING DATE: 2002-05-24
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-604A-7

Query Match      100.0%; Score 3103; DB 12; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.le-270;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFDLGEENFKALVLIAPAFYLOQCFFEDHVKLVNVEYFAKTCVADSAE 60
   |||||||
```

Db 25 DAHSEVAHRFKDGLGEENFKALVLIAPAFYLOQCPPEDEHVLVNEVTEFAKTCVADSAE 84  
QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPNLPRIVPEV 120  
Db 85 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPNLPRIVPEV 144  
QY 121 DVMTAFPHDNEETFLKKLYIETIARRHPYFYAPPELLFAKRYKAATTECCQAADKAACLLP 180  
Db 145 DVMTAFPHDNEETFLKKLYIETIARRHPYFYAPPELLFAKRYKAATTECCQAADKAACLLP 204  
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSQRPFAEFAEVS KLIVDTLTK 240  
Db 205 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSQRPFAEFAEVS KLIVDTLTK 264  
QY 241 VHTCCBGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCBGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVWLLLRKATYETTLTK 360  
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVWLLLRKATYETTLTK 384  
QY 361 CAADPHECVAKYVDFEFPKPLVEEPONLIKONCELFOLGEYKFNALLVRYTKVQVST 420  
Db 385 CAADPHECVAKYVDFEFPKPLVEEPONLIKONCELFOLGEYKFNALLVRYTKVQVST 444  
QY 421 PTLVEVSRLNGKVGSKCKKHPKAMPKPCABDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480  
Db 445 PTLVEVSRLNGKVGSKCKKHPKAMPKPCABDYLSVVLNOLCVLHEKTPVSDRVTKCTES 504  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVKKPKAT 540  
Db 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVKKPKAT 564  
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585  
Db 565 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 609

RESULT 10  
US-10-153-064-7  
; Sequence 7, Application US/10153064  
; Publication No. US2002014281441  
; GENERAL INFORMATION:  
; APPLICANT: Bell et al.  
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
; FILE REFERENCE: PF556  
; CURRENT APPLICATION NUMBER: US/10/153,064  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,212  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-153-064-7

Query Match 100.0%; Score 3103; DB 14; Length 609;  
Best Local Similarity 100.0%; Pred. No. 1.1e-270;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGLGEENFKALVLIAPAFYLOQCPPEDEHVLVNEVTEFAKTCVADSAE 60  
Db 25 DAHSEVAHRFKDGLGEENFKALVLIAPAFYLOQCPPEDEHVLVNEVTEFAKTCVADSAE 84  
QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPNLPRIVPEV 120  
Db 85 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPNLPRIVPEV 144  
QY 121 DVMTAFPHDNEETFLKKLYIETIARRHPYFYAPPELLFAKRYKAATTECCQAADKAACLLP 180

Db 145 DVMTAFPHDNEETFLKKLYIETIARRHPYFYAPPELLFAKRYKAATTECCQAADKAACLLP 204  
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSQRPFAEFAEVS KLIVDTLTK 240  
Db 205 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSQRPFAEFAEVS KLIVDTLTK 264  
QY 241 VHTCCBGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
Db 265 VHTCCBGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVWLLLRKATYETTLTK 360  
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVWLLLRKATYETTLTK 384  
QY 361 CAADPHECVAKYVDFEFPKPLVEEPONLIKONCELFOLGEYKFNALLVRYTKVQVST 420  
Db 385 CAADPHECVAKYVDFEFPKPLVEEPONLIKONCELFOLGEYKFNALLVRYTKVQVST 444  
QY 421 PTLVEVSRLNGKVGSKCKKHPKAMPKPCABDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480  
Db 445 PTLVEVSRLNGKVGSKCKKHPKAMPKPCABDYLSVVLNOLCVLHEKTPVSDRVTKCTES 504  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVKKPKAT 540  
Db 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVKKPKAT 564  
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585  
Db 565 KEQLKAVMDDDFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 609

RESULT 11  
US-09-984-186-2  
; Sequence 2, Application US/09984186  
; Patent No. US20020151011A1  
; GENERAL INFORMATION:  
; APPLICANT: Fieer, Reinhard  
; Fournier, Alain  
; Guittion, Jean-Dominique  
; Jung, Gerard  
; Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3043  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
; COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (PatentIn)  
; CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/984,186  
FILING DATE: 29-Oct-2001  
CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619



REFERENCE/DOCKET NUMBER: S92006-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 610 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-984-186-2

Query Match 100.0%; Score 3103; DB 10; Length 610;  
Best Local Similarity 100.0%; Pred. No. 1.1e-270;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDGBENFKALVIAFAQYLQCCPEDHVKLVNVEVTEFAKTCVADESAE 60  
DB 25 DAHSEVAHFKDGBENFKALVIAFAQYLQCCPEDHVKLVNVEVTEFAKTCVADESAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLVRPEV 144  
QY 121 DVMCTAFHDNEETFLKYLIEIARRHPYFYAPPELLFAKRYKAATTECCQAADKAACLLP 180  
DB 145 DVMCTAFHDNEETFLKYLIEIARRHPYFYAPPELLFAKRYKAATTECCQAADKAACLLP 204  
QY 181 KLDELDEGKASSAKORLKASLQKGFGERAFKAWAVARLSQRFPAEFAEVSILVTDLT 240  
DB 205 KLDELDEGKASSAKORLKASLQKGFGERAFKAWAVARLSQRFPAEFAEVSILVTDLT 264  
QY 241 VHTCCCHGDLLECADRADLAKYICENODSISKKLKECCPEKPLEKSHCIAEVENDEMPA 300  
DB 265 VHTCCCHGDLLECADRADLAKYICENODSISKKLKECCPEKPLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNTAEAKDVFGLMFLYEYARRHPDYSVLLRLAKRTYETILEK 360  
DB 325 DLPSLAADFVESKDVCKNTAEAKDVFGLMFLYEYARRHPDYSVLLRLAKRTYETILEK 384  
QY 361 CAADAPHECYAKVDFEFPKPLVEEPQNLTKQNCSELFQOLGEYKFQNALVRYTKKPVST 420  
DB 385 CAADAPHECYAKVDFEFPKPLVEEPQNLTKQNCSELFQOLGEYKFQNALVRYTKKPVST 444  
QY 421 PLYEVSRLNGKYGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 445 PLYEVSRLNGKYGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDYTPKEFNAETFTFHADICTLSEKROIKKOTALVELVKKPKAT 540  
DB 505 LVNRRPCFSALEVDYTPKEFNAETFTFHADICTLSEKROIKKOTALVELVKKPKAT 564  
QY 541 KEQLKAVMDDFAAVFVEKCKADDKETCFAPEGKKLVAAASQAALGL 585  
DB 565 KEQLKAVMDDFAAVFVEKCKADDKETCFAPEGKKLVAAASQAALGL 609

RESULT 12  
US-10-237-667-2  
Sequence 2, Application US/10237667  
Publication No. US20030022308A1  
GENERAL INFORMATION:  
APPLICANT: Fleer, Reinhard  
Fournier, Alain  
Guittion, Jean-Dominique  
Jung, Gerard  
Jeh, Patrice  
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
CONTAINING SAID POLYPEPTIDES  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (patentin)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/237,667  
FILING DATE: 10-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: S92006-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 610 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-237-667-2

Query Match 100.0%; Score 3103; DB 15; Length 610;  
Best Local Similarity 100.0%; Pred. No. 1.1e-270;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDGBENFKALVIAFAQYLQCCPEDHVKLVNVEVTEFAKTCVADESAE 60  
DB 25 DAHSEVAHFKDGBENFKALVIAFAQYLQCCPEDHVKLVNVEVTEFAKTCVADESAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLVRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLVRPEV 144  
QY 121 DVMCTAFHDNEETFLKYLIEIARRHPYFYAPPELLFAKRYKAATTECCQAADKAACLLP 180  
DB 145 DVMCTAFHDNEETFLKYLIEIARRHPYFYAPPELLFAKRYKAATTECCQAADKAACLLP 204  
QY 181 KLDELDEGKASSAKORLKASLQKGFGERAFKAWAVARLSQRFPAEFAEVSILVTDLT 240  
DB 205 KLDELDEGKASSAKORLKASLQKGFGERAFKAWAVARLSQRFPAEFAEVSILVTDLT 264  
QY 241 VHTCCCHGDLLECADRADLAKYICENODSISKKLKECCPEKPLEKSHCIAEVENDEMPA 300  
DB 265 VHTCCCHGDLLECADRADLAKYICENODSISKKLKECCPEKPLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNTAEAKDVFGLMFLYEYARRHPDYSVLLRLAKRTYETILEK 360  
DB 325 DLPSLAADFVESKDVCKNTAEAKDVFGLMFLYEYARRHPDYSVLLRLAKRTYETILEK 384  
QY 361 CAADAPHECYAKVDFEFPKPLVEEPQNLTKQNCSELFQOLGEYKFQNALVRYTKKPVST 420  
DB 385 CAADAPHECYAKVDFEFPKPLVEEPQNLTKQNCSELFQOLGEYKFQNALVRYTKKPVST 444  
QY 421 PLYEVSRLNGKYGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

Db 445 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVNLQCVLHKTVPVSDRVTKCTES 504  
QY 481 LVNRRPCFSALVEDETVYVPKFNATFTFHADICTLSEKEROIKKOTALVELVHKPKAT 540  
Db 505 LVNRRPCFSALVEDETVYVPKFNATFTFHADICTLSEKEROIKKOTALVELVHKPKAT 564  
QY 541 KEQLKAVMDFFAAFEVKCKADDDKTCFAEBGKKLVAASQAALGL 585  
Db 565 KEQLKAVMDFFAAFEVKCKADDDKTCFAEBGKKLVAASQAALGL 609

RESULT 13  
US-10-237-708-2  
; Sequence 2, Application US/10237708  
; Publication No. US20030036170A1  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; Fournier, Alain  
; Guitton, Jean-Dominique  
; Jung, Gerard  
; Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.1 (PatentIn)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/237,708  
; FILING DATE: 10-Sep-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/797,689  
; FILING DATE: 31-JAN-1997  
; APPLICATION NUMBER: US 08/256,927  
; FILING DATE: 28-JUL-1994  
; APPLICATION NUMBER: FR 92/01064  
; FILING DATE: 31-JAN-1992  
; APPLICATION NUMBER: PCT/FR93/00085  
; FILING DATE: 28-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith Ph.D., Julie K.  
; REGISTRATION NUMBER: P-39,619  
; REFERENCE/DOCKET NUMBER: S192006-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3839  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-237-708-2  
Query Match 100.0%; Score 3103; DB 15; Length 610;  
Best Local Similarity 100.0%; Pred. No. 1.1e-270;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHFRKDLGSENFKALVLIATFAQYILQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
|||||

Db 25 DAHSEVAHFRKDLGSENFKALVLIATFAQYILQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84  
QY 61 NCDLSLHTLFGDKLCVATLRETYGEMADCCAKOPEPERNECFLOHKDDNPNLPLRVPEY 120  
Db 85 NCDLSLHTLFGDKLCVATLRETYGEMADCCAKOPEPERNECFLOHKDDNPNLPLRVPEY 144  
QY 121 DVMCTAFHDNEETFLKKLYLIEIARHPYFYAPELLEFFAKRYKAAFTFECQQAADKAACILP 180  
Db 145 DVMCTAFHDNEETFLKKLYLIEIARHPYFYAPELLEFFAKRYKAAFTFECQQAADKAACILP 204  
QY 181 KIDELRDEGKASSAKORLKCASLOKFGRAFKAWAVARLSORPPKAEFAEYSKLYTDLTK 240  
Db 205 KIDELRDEGKASSAKORLKCASLOKFGRAFKAWAVARLSORPPKAEFAEYSKLYTDLTK 264  
QY 241 VTECHGDLLECADRADLAKYICENODSISKKLKECKEPLLEKSHCTIAEVENDEMPA 300  
Db 265 VTECHGDLLECADRADLAKYICENODSISKKLKECKEPLLEKSHCTIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARHPDYSVLLLRILAKTYETLEK 360  
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARHPDYSVLLLRILAKTYETLEK 384  
QY 361 CAADPHECHAKYVDFEFKPLVEEPONLLKONCELFEOIGYKFNALLVRYTKKVPQVST 420  
Db 385 CAADPHECHAKYVDFEFKPLVEEPONLLKONCELFEOIGYKFNALLVRYTKKVPQVST 444  
QY 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVNLQCVLHKTVPVSDRVTKCTES 480  
Db 445 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVNLQCVLHKTVPVSDRVTKCTES 504  
QY 481 LVNRRPCFSALVEDETVYVPKFNATFTFHADICTLSEKEROIKKOTALVELVHKPKAT 540  
Db 505 LVNRRPCFSALVEDETVYVPKFNATFTFHADICTLSEKEROIKKOTALVELVHKPKAT 564  
QY 541 KEQLKAVMDFFAAFEVKCKADDDKTCFAEBGKKLVAASQAALGL 585  
Db 565 KEQLKAVMDFFAAFEVKCKADDDKTCFAEBGKKLVAASQAALGL 609

RESULT 14  
US-10-237-866-2  
; Sequence 2, Application US/10237866  
; Publication No. US20030036171A1  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; Fournier, Alain  
; Guitton, Jean-Dominique  
; Jung, Gerard  
; Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.1 (PatentIn)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/237,866  
; FILING DATE: 10-Sep-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/797,689  
; FILING DATE: 31-JAN-1997  
; APPLICATION NUMBER: US 08/256,927  
; FILING DATE: 28-JUL-1994  
; APPLICATION NUMBER: FR 92/01064  
; FILING DATE: 31-JAN-1992  
; APPLICATION NUMBER: PCT/FR93/00085  
; FILING DATE: 28-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith Ph.D., Julie K.  
; REGISTRATION NUMBER: P-39,619  
; REFERENCE/DOCKET NUMBER: S192006-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3839  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-237-866-2  
Query Match 100.0%; Score 3103; DB 15; Length 610;  
Best Local Similarity 100.0%; Pred. No. 1.1e-270;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHSEVAHFRKDLGSENFKALVLIATFAQYILQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
|||||

```

;
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-866-2

Query Match 100.0%; Score 3103; DB 15; Length 610;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIAPAFYLOQCPEDHVKLVNVEVTEFAKTCVADESAE 60
Db 25 DAHSEVAHFRKDLGEENFKALVLIAPAFYLOQCPEDHVKLVNVEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPLRLVPRV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPLRLVPRV 144
QY 121 DVMCTAFPHDNEETFLKKYLYEIARRHPYFYAPPELLFAKRYKAAPTECCQAADKAACLLP 180
Db 145 DVMCTAFPHDNEETFLKKYLYEIARRHPYFYAPPELLFAKRYKAAPTECCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKCSLOKGERAFKAWARLSQRFPAKFAEVSKLIVDLTK 240
Db 205 KLDELDEGKASSAKQRLKCSLOKGERAFKAWARLSQRFPAKFAEVSKLIVDLTK 264
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSKHSCIAEVENDEMPA 300
Db 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSKHSCIAEVENDEMPA 324
QY 301 DLPISLAADFVESKDVCKNYAEAKDVFGLGMFLYIYARRHPDYSVLLLLSLAKTYETLEKC 360
Db 325 DLPISLAADFVESKDVCKNYAEAKDVFGLGMFLYIYARRHPDYSVLLLLSLAKTYETLEKC 384
QY 361 CAADPHECYAKVDFDEKPLVEEPONLIKONCELFQGLGEYKFNALLVRYTKKVPYST 420
Db 385 CAADPHECYAKVDFDEKPLVEEPONLIKONCELFQGLGEYKFNALLVRYTKKVPYST 444
QY 421 PTLVEVRNKGKVGSKCKHPKAPCAEDYLSVLNQLCVLHKKTPVSDRVTKCTES 480
Db 445 PTLVEVRNKGKVGSKCKHPKAPCAEDYLSVLNQLCVLHKKTPVSDRVTKCTES 504
QY 481 LVNRRPFCFSALEVDYVYKFNAAETFFHADICTLSEKEROIKKQTFALVELVKKPKAT 540
Db 505 LVNRRPFCFSALEVDYVYKFNAAETFFHADICTLSEKEROIKKQTFALVELVKKPKAT 564
QY 541 KEQLKAYMDFAAFVEKCKADDDKTCFPAEKGKKLVAASQAALGL 585
Db 565 KEQLKAYMDFAAFVEKCKADDDKTCFPAEKGKKLVAASQAALGL 609

RESULT 15
US-10-237-871-2
; Sequence 2, Application US/10237871
; Publication No. US20030036172A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain

```

```

;
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3c43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,871
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-871-2

Query Match 100.0%; Score 3103; DB 15; Length 610;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIAPAFYLOQCPEDHVKLVNVEVTEFAKTCVADESAE 60
Db 25 DAHSEVAHFRKDLGEENFKALVLIAPAFYLOQCPEDHVKLVNVEVTEFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPLRLVPRV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPLRLVPRV 144
QY 121 DVMCTAFPHDNEETFLKKYLYEIARRHPYFYAPPELLFAKRYKAAPTECCQAADKAACLLP 180
Db 145 DVMCTAFPHDNEETFLKKYLYEIARRHPYFYAPPELLFAKRYKAAPTECCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKCSLOKGERAFKAWARLSQRFPAKFAEVSKLIVDLTK 240
Db 205 KLDELDEGKASSAKQRLKCSLOKGERAFKAWARLSQRFPAKFAEVSKLIVDLTK 264
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSKHSCIAEVENDEMPA 300
Db 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSKHSCIAEVENDEMPA 324
QY 301 DLPISLAADFVESKDVCKNYAEAKDVFGLGMFLYIYARRHPDYSVLLLLSLAKTYETLEKC 360

```

Db		325	DLPSLAADFVESKDVCKNVAEAKDVFLGMLFYIYARRHPDYSVLLLELAKYTTLEKC	384
QY		361	CAAADPHECYAKVDFEFPQNLKONCELFQOLGEYKFQNALLVRYTKKVPQVST	420
Db		385	CAAADPHECYAKVDFEFPQNLKONCELFQOLGEYKFQNALLVRYTKKVPQVST	444
QY		421	PTLVEVSRNLGKVGSKCKHPEAKMPCARDYLSVVNLQLCVLHEKTPVSDRVTKCCTES	480
Db		445	PTLVEVSRNLGKVGSKCKHPEAKMPCARDYLSVVNLQLCVLHEKTPVSDRVTKCCTES	504
QY		481	LVNRRPCFSALEVEDEYVPKEFNATFTFHADICTLSEKEROIKKQOTALVELVYKHKPKAT	540
Db		505	LVNRRPCFSALEVEDEYVPKEFNATFTFHADICTLSEKEROIKKQOTALVELVYKHKPKAT	564
QY		541	KEQLKAYMDDFAAFVEKCKCKADDDKTCFABEGKLVAAQAALGL	585
Db		565	KEQLKAYMDDFAAFVEKCKCKADDDKTCFABEGKLVAAQAALGL	609

Search completed: August 31, 2003, 16:39:14  
 Job time : 30 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2003, 16:33:10 ; Search time 43 Seconds  
(without alignments)  
1308.341 Million cell updates/sec

Title: US-09-833-041-18  
Perfect score: 3103  
Sequence: 1 DAHKSEVAHFRKDLGEENFK.....TCFAEGKKLVLAASQAALGL 585

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	100.0	609	1 ABHUS	serum albumin prec
2	2942	94.8	600	2 A47391	serum albumin prec
3	2620	84.4	608	2 S57632	serum albumin prec
4	2475.5	79.8	607	1 ABHOS	serum albumin prec
5	2446.5	78.8	607	1 ABHOS	serum albumin prec
6	2432.5	78.4	607	1 ABHOS	serum albumin prec
7	2426	78.2	608	1 ABPTS	serum albumin prec
8	2411.5	77.7	605	1 ABPGS	serum albumin prec
9	2387	76.9	609	2 JC5838	albumin - Mongolia
10	1861	60.0	453	2 A05139	serum albumin - mo
11	1557.5	50.2	615	1 ABCHS	serum albumin prec
12	1253.5	40.4	609	2 JC4258	alpha-fetoprotein
13	1249.5	40.3	609	1 FEPU	alpha-fetoprotein
14	1242.5	40.0	609	1 FEPU	alpha-fetoprotein
15	1205	38.8	607	1 ABXL72	74K albumin precursor
16	1181.5	38.1	265	2 I46986	albumin - dog (fra
17	1175.5	37.9	608	1 ABX168	68K serum albumin
18	1084	34.9	605	1 FPMS	alpha-fetoprotein
19	1067	34.4	611	1 FPPT	alpha-fetoprotein
20	1055	34.0	599	1 A54906	afamin precursor -
21	928.5	29.9	614	2 S59517	serum albumin prec
22	928	29.9	608	2 A53195	afamin precursor -
23	747.5	24.1	608	1 ABONS1	serum albumin 1 pr
24	742.5	23.9	608	1 ABONS2	serum albumin 2 pr
25	699	22.5	382	2 A37253	serum albumin - bu
26	440.5	14.2	1423	1 S27941	serum albumin - se
27	386	12.4	474	1 VYHDD	vitamin D-binding
28	385	12.4	476	1 VYRTD	vitamin D-binding
29	372	12.0	472	1 A35327	vitamin D-binding

30 184 5.9 1819 2 A71928 cag island protein  
31 184 5.9 1927 2 G64585 cag pathogenicity  
32 141.5 4.6 1560 2 T30282 calcium-binding pr  
33 137.5 4.4 1348 2 AG2558 hypothetical prote  
34 134.5 4.3 1004 2 JC2221 major surface glyco  
35 134 4.3 1083 2 JC2300 cell surface glyco  
36 134 4.3 1780 2 T17272 hypothetical prote  
37 133.5 4.3 1390 2 S51364 sperm tail-specifi  
38 132.5 4.3 1070 2 T06733 kinesin homolog F2  
39 131 4.2 1076 2 JC2217 major surface glyco  
40 130 4.2 1175 2 D35815 myosin heavy chain  
41 129 4.2 1017 2 JC4035 cell-cycle-depende  
42 128.5 4.1 1051 2 PC4091 glycoprotein A - p  
43 128 4.1 1005 2 A64465 hypothetical prote  
44 128 4.1 3225 2 I52300 giantin - human  
45 126.5 4.1 1927 2 A59236 embryonic muscle m

ALIGNMENTS

RESULT 1

ABHUS  
serum albumin precursor [validated] - human  
N:Alternate names: preproalbumin  
N:Contains: kinetensin  
C:Species: Homo sapiens (man)  
C:Date: 29-Jul-1981 #sequence\_revision 31-Jan-1997 #text\_change 17-Mar-2000  
C:Accession: A93743; A93936; I39427; I59286; I59313; G01747; S5314; A91420; S06422;  
R:Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, C.M.; Houch, C.M.; Najarian, R.C.; Se  
Nucleic Acids Res. 9, 6103-6114, 1981  
A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia  
A:Reference number: A93743; MUID:82081882; PMID:6171778  
A:Accession: A93743  
A:Molecule type: mRNA  
A:Residues: 1-419, 'K', 421-609 <LAW>  
A:Cross-references: EMBL:Y00494; GB:J00078; GB:L00132; GB:L00133; NID:g28591; PIDN:C  
R:Dugaiczkyk, A.; Law, S.W.; Dennison, O.E.  
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982  
A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA  
A:Reference number: A93936; MUID:82105994; PMID:6273391  
A:Accession: A93936  
A:Molecule type: mRNA  
A:Residues: 1-120, 'G', 122-609 <DUG>  
A:Cross-references: EMBL:Y00494; NID:g28589; PIDN:CRAA23753.1; PID:g28590  
R:Grano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.  
J. Biol. Chem. 261, 3244-3251, 1986  
A:Title: The human albumin gene. Characterization of the 5' and 3' flanking regions  
A:Reference number: I39427; MUID:86140099; PMID:2419329  
A:Accession: I39427  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-26 <URA>  
A:Cross-references: GB:M13075; NID:g178330; PIDN:AAA51688.1; PID:g553173  
R:Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.  
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994  
A:Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian fam  
A:Reference number: I59286; MUID:94181575; PMID:8134387  
A:Accession: I59286  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 282-290, 'KSRFDLO' <WAT>  
A:Cross-references: GB:S69192; NID:g546032; PIDN:AAB30282.1; PID:g546033  
A:Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia  
R:Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putn  
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994  
A:Title: Genetic variants of human serum albumin in Italy: point mutants and a carb  
A:Reference number: I59313; MUID:94294404; PMID:8022807  
A:Accession: I59313  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 589-590, 'ALPRVKNLLQVLP' <MAD>  
A:Cross-references: GB:S70799; NID:g547231; PIDN:AAB31177.1; PID:g547232

- A;Note: this frame-shift variant is designated albumin Bazzano; four additional variants
- R;Menaya, J.; Parrilla, R.; Ayuso, M.S.  
submitted to the EMBL Data Library, March 1995
- A;Reference number: G08292
- A;Accession: G01747
- A;Status: translated from GB/EMBL/DBJ
- A;Molecule type: mRNA
- A;Residues: 1-122, 'G', 122-455 <MEN>
- A;Cross-references: EMBL:U22961; NID:G763428; PID:AAA64922.1; PID:G763431
- R;Leiglerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.  
Biochem. J. 308, 321-325, 1995
- A;Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast *Kex2*
- A;Reference number: S55314; MUID:95275251; PMID:7755381
- A;Accession: S55314
- A;Molecule type: protein
- A;Residues: 19-27 <LESD>
- R;Meloun, B.; Moravsek, L.; Kostka, V.  
FEBS Lett. 58, 134-137, 1975
- A;Title: Complete amino acid sequence of human serum albumin.
- A;Reference number: A91420; MUID:76187907; PMID:1225573
- A;Accession: A91420
- A;Molecule type: protein
- A;Residues: 25-117, 'Q', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395-396, 'D', 397-400, 'D', 401-402, 'D', 403-404, 'D', 405-406, 'D', 407-408, 'D', 409-410, 'D', 411-412, 'D', 413-414, 'D', 415-416, 'D', 417-418, 'D', 419-420, 'D', 421-422, 'D', 423-424, 'D', 425-426, 'D', 427-428, 'D', 429-430, 'D', 431-432, 'D', 433-434, 'D', 435-436, 'D', 437-438, 'D', 439-440, 'D', 441-442, 'D', 443-444, 'D', 445-446, 'D', 447-448, 'D', 449-450, 'D', 451-452, 'D', 453-454, 'D', 455-456, 'D', 457-458, 'D', 459-460, 'D', 461-462, 'D', 463-464, 'D', 465-466, 'D', 467-468, 'D', 469-470, 'D', 471-472, 'D', 473-474, 'D', 475-476, 'D', 477-478, 'D', 479-480, 'D', 481-482, 'D', 483-484, 'D', 485-486, 'D', 487-488, 'D', 489-490, 'D', 491-492, 'D', 493-494, 'D', 495-496, 'D', 497-498, 'D', 499-500, 'D', 501-502, 'D', 503-504, 'D', 505-506, 'D', 507-508, 'D', 509-510, 'D', 511-512, 'D', 513-514, 'D', 515-516, 'D', 517-518, 'D', 519-520, 'D', 521-522, 'D', 523-524, 'D', 525-526, 'D', 527-528, 'D', 529-530, 'D', 531-532, 'D', 533-534, 'D', 535-536, 'D', 537-538, 'D', 539-540, 'D', 541-542, 'D', 543-544, 'D', 545-546, 'D', 547-548, 'D', 549-550, 'D', 551-552, 'D', 553-554, 'D', 555-556, 'D', 557-558, 'D', 559-560, 'D', 561-562, 'D', 563-564, 'D', 565-566, 'D', 567-568, 'D', 569-570, 'D', 571-572, 'D', 573-574, 'D', 575-576, 'D', 577-578, 'D', 579-580, 'D', 581-582, 'D', 583-584, 'D', 585-586, 'D', 587-588, 'D', 589-590, 'D', 591-592, 'D', 593-594, 'D', 595-596, 'D', 597-598, 'D', 599-600, 'D', 601-602, 'D', 603-604, 'D', 605-606, 'D', 607-608, 'D', 609-610, 'D', 611-612, 'D', 613-614, 'D', 615-616, 'D', 617-618, 'D', 619-620, 'D', 621-622, 'D', 623-624, 'D', 625-626, 'D', 627-628, 'D', 629-630, 'D', 631-632, 'D', 633-634, 'D', 635-636, 'D', 637-638, 'D', 639-640, 'D', 641-642, 'D', 643-644, 'D', 645-646, 'D', 647-648, 'D', 649-650, 'D', 651-652, 'D', 653-654, 'D', 655-656, 'D', 657-658, 'D', 659-660, 'D', 661-662, 'D', 663-664, 'D', 665-666, 'D', 667-668, 'D', 669-670, 'D', 671-672, 'D', 673-674, 'D', 675-676, 'D', 677-678, 'D', 679-680, 'D', 681-682, 'D', 683-684, 'D', 685-686, 'D', 687-688, 'D', 689-690, 'D', 691-692, 'D', 693-694, 'D', 695-696, 'D', 697-698, 'D', 699-700, 'D', 701-702, 'D', 703-704, 'D', 705-706, 'D', 707-708, 'D', 709-710, 'D', 711-712, 'D', 713-714, 'D', 715-716, 'D', 717-718, 'D', 719-720, 'D', 721-722, 'D', 723-724, 'D', 725-726, 'D', 727-728, 'D', 729-730, 'D', 731-732, 'D', 733-734, 'D', 735-736, 'D', 737-738, 'D', 739-740, 'D', 741-742, 'D', 743-744, 'D', 745-746, 'D', 747-748, 'D', 749-750, 'D', 751-752, 'D', 753-754, 'D', 755-756, 'D', 757-758, 'D', 759-760, 'D', 761-762, 'D', 763-764, 'D', 765-766, 'D', 767-768, 'D', 769-770, 'D', 771-772, 'D', 773-774, 'D', 775-776, 'D', 777-778, 'D', 779-780, 'D', 781-782, 'D', 783-784, 'D', 785-786, 'D', 787-788, 'D', 789-790, 'D', 791-792, 'D', 793-794, 'D', 795-796, 'D', 797-798, 'D', 799-800, 'D', 801-802, 'D', 803-804, 'D', 805-806, 'D', 807-808, 'D', 809-810, 'D', 811-812, 'D', 813-814, 'D', 815-816, 'D', 817-818, 'D', 819-820, 'D', 821-822, 'D', 823-824, 'D', 825-826, 'D', 827-828, 'D', 829-830, 'D', 831-832, 'D', 833-834, 'D', 835-836, 'D', 837-838, 'D', 839-840, 'D', 841-842, 'D', 843-844, 'D', 845-846, 'D', 847-848, 'D', 849-850, 'D', 851-852, 'D', 853-854, 'D', 855-856, 'D', 857-858, 'D', 859-860, 'D', 861-862, 'D', 863-864, 'D', 865-866, 'D', 867-868, 'D', 869-870, 'D', 871-872, 'D', 873-874, 'D', 875-876, 'D', 877-878, 'D', 879-880, 'D', 881-882, 'D', 883-884, 'D', 885-886, 'D', 887-888, 'D', 889-890, 'D', 891-892, 'D', 893-894, 'D', 895-896, 'D', 897-898, 'D', 899-900, 'D', 901-902, 'D', 903-904, 'D', 905-906, 'D', 907-908, 'D', 909-910, 'D', 911-912, 'D', 913-914, 'D', 915-916, 'D', 917-918, 'D', 919-920, 'D', 921-922, 'D', 923-924, 'D', 925-926, 'D', 927-928, 'D', 929-930, 'D', 931-932, 'D', 933-934, 'D', 935-936, 'D', 937-938, 'D', 939-940, 'D', 941-942, 'D', 943-944, 'D', 945-946, 'D', 947-948, 'D', 949-950, 'D', 951-952, 'D', 953-954, 'D', 955-956, 'D', 957-958, 'D', 959-960, 'D', 961-962, 'D', 963-964, 'D', 965-966, 'D', 967-968, 'D', 969-970, 'D', 971-972, 'D', 973-974, 'D', 975-976, 'D', 977-978, 'D', 979-980, 'D', 981-982, 'D', 983-984, 'D', 985-986, 'D', 987-988, 'D', 989-990, 'D', 991-992, 'D', 993-994, 'D', 995-996, 'D', 997-998, 'D', 999-1000, 'D', 1001-1002, 'D', 1003-1004, 'D', 1005-1006, 'D', 1007-1008, 'D', 1009-1010, 'D', 1011-1012, 'D', 1013-1014, 'D', 1015-1016, 'D', 1017-1018, 'D', 1019-1020, 'D', 1021-1022, 'D', 1023-1024, 'D', 1025-1026, 'D', 1027-1028, 'D', 1029-1030, 'D', 1031-1032, 'D', 1033-1034, 'D', 1035-1036, 'D', 1037-1038, 'D', 1039-1040, 'D', 1041-1042, 'D', 1043-1044, 'D', 1045-1046, 'D', 1047-1048, 'D', 1049-1050, 'D', 1051-1052, 'D', 1053-1054, 'D', 1055-1056, 'D', 1057-1058, 'D', 1059-1060, 'D', 1061-1062, 'D', 1063-1064, 'D', 1065-1066, 'D', 1067-1068, 'D', 1069-1070, 'D', 1071-1072, 'D', 1073-1074, 'D', 1075-1076, 'D', 1077-1078, 'D', 1079-1080, 'D', 1081-1082, 'D', 1083-1084, 'D', 1085-1086, 'D', 1087-1088, 'D', 1089-1090, 'D', 1091-1092, 'D', 1093-1094, 'D', 1095-1096, 'D', 1097-1098, 'D', 1099-1100, 'D', 1101-1102, 'D', 1103-1104, 'D', 1105-1106, 'D', 1107-1108, 'D', 1109-1110, 'D', 1111-1112, 'D', 1113-1114, 'D', 1115-1116, 'D', 1117-1118, 'D', 1119-1120, 'D', 1121-1122, 'D', 1123-1124, 'D', 1125-1126, 'D', 1127-1128, 'D', 1129-1130, 'D', 1131-1132, 'D', 1133-1134, 'D', 1135-1136, 'D', 1137-1138, 'D', 1139-1140, 'D', 1141-1142, 'D', 1143-1144, 'D', 1145-1146, 'D', 1147-1148, 'D', 1149-1150, 'D', 1151-1152, 'D', 1153-1154, 'D', 1155-1156, 'D', 1157-1158, 'D', 1159-1160, 'D', 1161-1162, 'D', 1163-1164, 'D', 1165-1166, 'D', 1167-1168, 'D', 1169-1170, 'D', 1171-1172, 'D', 1173-1174, 'D', 1175-1176, 'D', 1177-1178, 'D', 1179-1180, 'D', 1181-1182, 'D', 1183-1184, 'D', 1185-1186, 'D', 1187-1188, 'D', 1189-1190, 'D', 1191-1192, 'D', 1193-1194, 'D', 1195-1196, 'D', 1197-1198, 'D', 1199-1200, 'D', 1201-1202, 'D', 1203-1204, 'D', 1205-1206, 'D', 1207-1208, 'D', 1209-1210, 'D', 1211-1212, 'D', 1213-1214, 'D', 1215-1216, 'D', 1217-1218, 'D', 1219-1220, 'D', 1221-1222, 'D', 1223-1224, 'D', 1225-1226, 'D', 1227-1228, 'D', 1229-1230, 'D', 1231-1232, 'D', 1233-1234, 'D', 1235-1236, 'D', 1237-1238, 'D', 1239-1240, 'D', 1241-1242, 'D', 1243-1244, 'D', 1245-1246, 'D', 1247-1248, 'D', 1249-1250, 'D', 1251-1252, 'D', 1253-1254, 'D', 1255-1256, 'D', 1257-1258, 'D', 1259-1260, 'D', 1261-1262, 'D', 1263-1264, 'D', 1265-1266, 'D', 1267-1268, 'D', 1269-1270, 'D', 1271-1272, 'D', 1273-1274, 'D', 1275-1276, 'D', 1277-1278, 'D', 1279-1280, 'D', 1281-1282, 'D', 1283-1284, 'D', 1285-1286, 'D', 1287-1288, 'D', 1289-1290, 'D', 1291-1292, 'D', 1293-1294, 'D', 1295-1296, 'D', 1297-1298, 'D', 1299-1300, 'D', 1301-1302, 'D', 1303-1304, 'D', 1305-1306, 'D', 1307-1308, 'D', 1309-1310, 'D', 1311-1312, 'D', 1313-1314, 'D', 1315-1316, 'D', 1317-1318, 'D', 1319-1320, 'D', 1321-1322, 'D', 1323-1324, 'D', 1325-1326, 'D', 1327-1328, 'D', 1329-1330, 'D', 1331-1332, 'D', 1333-1334, 'D', 1335-1336, 'D', 1337-1338, 'D', 1339-1340, 'D', 1341-1342, 'D', 1343-1344, 'D', 1345-1346, 'D', 1347-1348, 'D', 1349-1350, 'D', 1351-1352, 'D', 1353-1354, 'D', 1355-1356, 'D', 1357-1358, 'D', 1359-1360, 'D', 1361-1362, 'D', 1363-1364, 'D', 1365-1366, 'D', 1367-1368, 'D', 1369-1370, 'D', 1371-1372, 'D', 1373-1374, 'D', 1375-1376, 'D', 1377-1378, 'D', 1379-1380, 'D', 1381-1382, 'D', 1383-1384, 'D', 1385-1386, 'D', 1387-1388, 'D', 1389-1390, 'D', 1391-1392, 'D', 1393-1394, 'D', 1395-1396, 'D', 1397-1398, 'D', 1399-1400, 'D', 1401-1402, 'D', 1403-1404, 'D', 1405-1406, 'D', 1407-1408, 'D', 1409-1410, 'D', 1411-1412, 'D', 1413-1414, 'D', 1415-1416, 'D', 1417-1418, 'D', 1419-1420, 'D', 1421-1422, 'D', 1423-1424, 'D', 1425-1426, 'D', 1427-1428, 'D', 1429-1430, 'D', 1431-1432, 'D', 1433-1434, 'D', 1435-1436, 'D', 1437-1438, 'D', 1439-1440, 'D', 1441-1442, 'D', 1443-1444, 'D', 1445-1446, 'D', 1447-1448, 'D', 1449-1450, 'D', 1451-1452, 'D', 1453-1454, 'D', 1455-1456, 'D', 1457-1458, 'D', 1459-1460, 'D', 1461-1462, 'D', 1463-1464, 'D', 1465-1466, 'D', 1467-1468, 'D', 1469-1470, 'D', 1471-1472, 'D', 1473-1474, 'D', 1475-1476, 'D', 1477-1478, 'D', 1479-1480, 'D', 1481-1482, 'D', 1483-1484, 'D', 1485-1486, 'D', 1487-1488, 'D', 1489-1490, 'D', 1491-1492, 'D', 1493-1494, 'D', 1495-1496, 'D', 1497-1498, 'D', 1499-1500, 'D', 1501-1502, 'D', 1503-1504, 'D', 1505-1506, 'D', 1507-1508, 'D', 1509-1510, 'D', 1511-1512, 'D', 1513-1514, 'D', 1515-1516, 'D', 1517-1518, 'D', 1519-1520, 'D', 1521-1522, 'D', 1523-1524, 'D', 1525-1526, 'D', 1527-1528, 'D', 1529-1530, 'D', 1531-1532, 'D', 1533-1534, 'D', 1535-1536, 'D', 1537-1538, 'D', 1539-1540, 'D', 1541-1542, 'D', 1543-1544, 'D', 1545-1546, 'D', 1547-1548, 'D', 1549-1550, 'D', 1551-1552, 'D', 1553-1554, 'D', 1555-1556, 'D', 1557-1558, 'D', 1559-1560, 'D', 1561-1562, 'D', 1563-1564, 'D', 1565-1566, 'D', 1567-1568, 'D', 1569-1570, 'D', 1571-1572, 'D', 1573-1574, 'D', 1575-1576, 'D', 1577-1578, 'D', 1579-1580, 'D', 1581-1582, 'D', 1583-1584, 'D', 1585-1586, 'D', 1587-1588, 'D', 1589-1590, 'D', 1591-1592, 'D', 1593-1594, 'D', 1595-1596, 'D', 1597-1598, 'D', 1599-1600, 'D', 1601-1602, 'D', 1603-1604, 'D', 1605-1606, 'D', 1607-1608, 'D', 1609-1610, 'D', 1611-1612, 'D', 1613-1614, 'D', 1615-1616, 'D', 1617-1618, 'D', 1619-1620, 'D', 1621-1622, 'D', 1623-1624, 'D', 1625-1626, 'D', 1627-1628, 'D', 1629-1630, 'D', 1631-1632, 'D', 1633-1634, 'D', 1635-1636, 'D', 1637-1638, 'D', 1639-1640, 'D', 1641-1642, 'D', 1643-1644, 'D', 1645-1646, 'D', 1647-1648, 'D', 1649-1650, 'D', 1651-1652, 'D', 1653-1654, 'D', 1655-1656, 'D', 1657-1658, 'D', 1659-1660, 'D', 1661-1662, 'D', 1663-1664, 'D', 1665-1666, 'D', 1667-1668, 'D', 1669-1670, 'D', 1671-1672, 'D', 1673-1674, 'D', 1675-1676, 'D', 1677-1678, 'D', 1679-1680, 'D', 1681-1682, 'D', 1683-1684, 'D', 1685-1686, 'D', 1687-1688, 'D', 1689-1690, 'D', 1691-1692, 'D', 1693-1694, 'D', 1695-1696, 'D', 1697-1698, 'D', 1699-1700, 'D', 1701-1702, 'D', 1703-1704, 'D', 1705-1706, 'D', 1707-1708, 'D', 1709-1710, 'D', 1711-1712, 'D', 1713-1714, 'D', 1715-1716, 'D', 1717-1718, 'D', 1719-1720, 'D', 1721-1722, 'D', 1723-1724, 'D', 1725-1726, 'D', 1727-1728, 'D', 1729-1730, 'D', 1731-1732, 'D', 1733-1734, 'D', 1735-1736, 'D', 1737-1738, 'D', 1739-1740, 'D', 1741-1742, 'D', 1743-1744, 'D', 1745-1746, 'D', 1747-1748, 'D', 1749-1750, 'D', 1751-1752, 'D', 1753-1754, 'D', 1755-1756, 'D', 1757-1758, 'D', 1759-1760, 'D', 1761-1762, 'D', 1763-1764, 'D', 1765-1766, 'D', 1767-1768, 'D', 1769-1770, 'D', 1771-1772, 'D', 1773-1774, 'D', 1775-1776, 'D', 1777-1778, 'D', 1779-1780, 'D', 1781-1782, 'D', 1783-1784, 'D', 1785-1786, 'D', 1787-1788, 'D', 1789-1790, 'D', 1791-1792, 'D', 1793-1794, 'D', 1795-1796, 'D', 1797-1798, 'D', 1799-1800, 'D', 1801-1802, 'D', 1803-1804, 'D', 1805-1806, 'D', 1807-1808, 'D', 1809-1810, 'D', 1811-1812, 'D', 1813-1814, 'D', 1815-1816, 'D', 1817-1818, 'D', 1819-1820, 'D', 1821-1822, 'D', 1823-1824, 'D', 1825-1826, 'D', 1827-1828, 'D', 1829-1830, 'D', 1831-1832, 'D', 1833-1834, 'D', 1835-1836, 'D', 1837-1838, 'D', 1839-1840, 'D', 1841-1842, 'D', 1843-1844, 'D', 1845-1846, 'D', 1847-1848, 'D', 1849-1850, 'D', 1851-1852, 'D', 1853-1854, 'D', 1855-1856, 'D', 1857-1858, 'D', 1859-1860, 'D', 1861-1862, 'D', 1863-1864, 'D', 1865-1866, 'D', 1867-1868, 'D', 1869-1870, 'D', 1871-1872, 'D', 1873-1874, 'D', 1875-1876, 'D', 1877-1878, 'D', 1879-1880, 'D', 1881-1882, 'D', 1883-1884, 'D', 1885-1886, 'D', 1887-1888, 'D', 1889-1890, 'D', 1891-1892, 'D', 1893-1894, 'D', 1895-1896, 'D', 1897-1898, 'D', 1899-1900, 'D', 1901-1902, 'D', 1903-1904, 'D', 1905-1906, 'D', 1907-1908, 'D', 1909-1910, 'D', 1911-1912, 'D', 1913-1914, 'D', 1915-1916, 'D', 1917-1918, 'D', 1919-1920, 'D', 1921-1922, 'D', 1923-1924, 'D', 1925-1926, 'D', 1927-1928, 'D', 1929-1930, 'D', 1931-1932, 'D', 1933-1934, 'D', 1935-1936, 'D', 1937-1938, 'D', 1939-1940, 'D', 1941-1942, 'D', 1943-1944, 'D', 1945-1946, 'D', 1947-1948, 'D', 1949-1950, 'D', 1951-1952, 'D', 1953-1954, 'D', 1955-1956, 'D', 1957-1958, 'D', 1959-1960, 'D', 1961-1962, 'D', 1963-1964, 'D', 1965-1966, 'D', 1967-1968, 'D', 1969-1970, 'D', 1971-1972, 'D', 1973-1974, 'D', 1975-1976, 'D', 1977-1978, 'D', 1979-1980, 'D', 1981-1982, 'D', 1983-1984, 'D', 1985-1986, 'D', 1987-1988, 'D', 1989-1990, 'D', 1991-1992, 'D', 1993-1994, 'D', 1995-1996, 'D', 1997-1998, 'D', 1999-2000, 'D', 2001-2002, 'D', 2003-2004, 'D', 2005-2006, 'D', 2007-2008, 'D', 2009-2010, 'D', 2011-2012, 'D', 2013-2014, 'D', 2015-2016, 'D', 2017-2018, 'D', 2019-2020, 'D', 2021-2022, 'D', 2023-2024, 'D', 2025-2026, 'D', 2027-2028, 'D', 2029-2030, 'D', 2031-2032, 'D', 2033-2034, 'D', 2035-2036, 'D', 2037-2038, 'D', 2039-2040, 'D', 2041-2042, 'D', 2043-2044, 'D', 2045-2046, 'D', 2047-2048, 'D', 2049-2050, 'D', 2051-2052, 'D', 2053-2054, 'D', 2055-2056, 'D', 2057-2058, 'D', 2059-2060, 'D', 2061-2062, 'D', 2063-2064, 'D', 2065-2066, 'D', 2067-2068, 'D', 2069-2070, 'D', 2071-2072, 'D', 2073-2074, 'D', 2075-2076, 'D', 2077-2078, 'D', 2079-2080, 'D', 2081-2082, 'D', 2083-2084, 'D', 2085-2086, 'D', 2087-2088, 'D', 2089-2090, 'D', 2091-2092, 'D', 2093-2094, 'D', 2095-2096, 'D', 2097-2098, 'D', 2099-2100, 'D', 2101-2102, 'D', 2103-2104, 'D', 2105-2106, 'D', 2107-2108, 'D', 2109-2110, 'D', 2111-2112, 'D', 2113-2114, 'D', 2115-2116, 'D', 2117-2118, 'D', 2119-2120, 'D', 2121-2122, 'D', 2123-2124, 'D', 2125-2126, 'D', 2127-2128, 'D', 2129-2130, 'D', 2131-2132, 'D', 2133-2134, 'D', 2135-2136, 'D', 2137-2138, 'D', 2139-2140, 'D', 2141-2142, 'D', 2143-21

F;166-174/Product: kinetensin #status experimental <KIP>  
F;221-394/Domain: serum albumin repeat homology <SA2>  
F;413-592/Domain: serum albumin repeat homology <SA2>  
F;27/Binding site: copper (His) #status predicted  
F;77-86,99-115,114-125,148-193,192-201,224-270,269-303,302-313,340-385,384-393,4  
F;214/Binding site: pyridoxal phosphate (lys) (covalent) #status experimental

Query Match 100.0%; Score 3103; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 9,2e-198;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDGLGENFKALVLIAPAFYLOQCFFEDHVKLVNVEFEAKTCVADESAR 60  
DB 25 DAHSEVAHFKDGLGENFKALVLIAPAFYLOQCFFEDHVKLVNVEFEAKTCVADESAR 84  
QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120  
DB 85 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 144  
QY 121 DVMCTAFHNEETFLAKYIETARHPHYFAPPELLFFAKRYKAATFCCQAADKAACLLP 180  
DB 145 DVMCTAFHNEETFLAKYIETARHPHYFAPPELLFFAKRYKAATFCCQAADKAACLLP 204  
QY 181 KLDELDEGKASSAKORLKASLOKFGRAFAKAWAVARLSORFPKAEFAVSKLVTDLT 240  
DB 205 KLDELDEGKASSAKORLKASLOKFGRAFAKAWAVARLSORFPKAEFAVSKLVTDLT 264  
QY 241 VHTCCGHDLLLECADRADLAKYICENODSISSKLKECCPELLEKSHCIAEVENDEMPA 300  
DB 265 VHTCCGHDLLLECADRADLAKYICENODSISSKLKECCPELLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLEYARRHPDYSVLLLRKATYETILEKC 360  
DB 325 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLEYARRHPDYSVLLLRKATYETILEKC 384  
QY 361 CAADPHECYAKVDFEFPLVEEPONLIKQNCLEFQGLGEYKFNALLVRYTKVPQVST 420  
DB 385 CAADPHECYAKVDFEFPLVEEPONLIKQNCLEFQGLGEYKFNALLVRYTKVPQVST 444  
QY 421 PTLVEVSRLNGKVGSKCKHPKAMPKPCADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRLNGKVGSKCKHPKAMPKPCADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504  
QY 481 LVNRPFCPSALEVDYTPVKFNATFTFHADICTLSEKROIKKOTALVELVKKPKAT 540  
DB 505 LVNRPFCPSALEVDYTPVKFNATFTFHADICTLSEKROIKKOTALVELVKKPKAT 564  
QY 541 KEQLKAVMDPFAFVEKCKCKADDKTCFABEGKLVAAASQAALGL 585  
DB 565 KEQLKAVMDPFAFVEKCKCKADDKTCFABEGKLVAAASQAALGL 609

RESULT 2  
A47391  
serum albumin precursor - rhesus macaque  
C:Species: Macaca mulatta (rhesus macaque)  
C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Aug-1999  
C:Accession: A47391  
R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dweulet, J.; Putnam, F.  
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993  
A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in biliru  
A:Reference number: A47391; MUID:93211971; PMID:8460152  
A:Contents: B/B homozygote  
A:Accession: A47391  
A:Status: preliminary  
A:Molecule type: mRNA; protein  
A:Residues: 1-600 <WAT>  
A:Cross-references: GB:M90463; NID:g342294; PIDN:AAA36906.1; PID:g342295  
A:Experimental source: liver  
A:Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBI:P:128281).  
C:Superfamily: serum albumin; serum albumin repeat homology  
F;21-194/Domain: serum albumin repeat homology <SA1>  
F;213-386/Domain: serum albumin repeat homology <SA2>

F;405-584/Domain: serum albumin repeat homology <SA3>  
Query Match 94.8%; Score 2942; DB 2; Length 600;  
Best Local Similarity 93.5%; Pred. No. 4e-187;  
Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDGLGENFKALVLIAPAFYLOQCFFEDHVKLVNVEFEAKTCVADESAR 60  
DB 17 DTHKSEVAHFKDGLGEEHGLVAFSFOYLOQCFFEEHVKLVNVEFEAKTCVADESAR 76  
QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120  
DB 77 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 136  
QY 121 DVMCTAFHNEETFLAKYIETARHPHYFAPPELLFFAKRYKAATFCCQAADKAACLLP 180  
DB 137 DVMCTAFHNEETFLAKYIETARHPHYFAPPELLFFAKRYKAATFCCQAADKAACLLP 196  
QY 181 KLDELDEGKASSAKORLKASLOKFGRAFAKAWAVARLSORFPKAEFAVSKLVTDLT 240  
DB 197 KLDELDEGKASSAKORLKASLOKFGRAFAKAWAVARLSORFPKAEFAVSKLVTDLT 256  
QY 241 VHTCCGHDLLLECADRADLAKYICENODSISSKLKECCPELLEKSHCIAEVENDEMPA 300  
DB 257 VHTCCGHDLLLECADRADLAKYICENODSISSKLKECCPELLEKSHCIAEVENDEMPA 316  
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLEYARRHPDYSVLLLRKATYETILEKC 360  
DB 317 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLEYARRHPDYSVLLLRKATYETILEKC 376  
QY 361 CAADPHECYAKVDFEFPLVEEPONLIKQNCLEFQGLGEYKFNALLVRYTKVPQVST 420  
DB 377 CAADPHECYAKVDFEFPLVEEPONLIKQNCLEFQGLGEYKFNALLVRYTKVPQVST 436  
QY 421 PTLVEVSRLNGKVGSKCKHPKAMPKPCADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 437 PTLVEVSRLNGKVGSKCKHPKAMPKPCADYLSVVLNQLCVLHEKTPVSEKVTTCCTES 496  
QY 481 LVNRPFCPSALEVDYTPVKFNATFTFHADICTLSEKROIKKOTALVELVKKPKAT 540  
DB 497 LVNRPFCPSALEVDYTPVKFNATFTFHADICTLSEKROIKKOTALVELVKKPKAT 556  
QY 541 KEQLKAVMDPFAFVEKCKCKADDKTCFABEGKLVAAASQAAL 583  
DB 557 KEQLKAVMDPFAFVEKCKCKADDKTCFABEGKLVAAASQAAL 599

RESULT 3  
S57632  
serum albumin precursor - cat  
C:Species: Felis silvestris catus (domestic cat)  
C:Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 20-Aug-1999  
C:Accession: J04660; S57632  
R:Hilger, C.; Grigioni, F.; Hentges, F.  
Gene 169, 295-296, 1996  
A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.  
A:Reference number: J04660; MUID:96194824; PMID:8647469  
A:Accession: J04660  
A:Molecule type: mRNA  
A:Residues: 1-608 <H12>  
A:Cross-references: EMBL:X84842; NID:g886484; PIDN:CAA59279.1; PID:g886485  
A:Experimental source: liver  
C:Comment: This protein is the major protein component in plasma. It functions as a  
ein has 35 conserved cysteine residues.  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: liver; plasma  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-24/Domain: propeptide #status predicted <PRP>  
F;25-608/Product: serum albumin #status predicted <MAT>  
F;29-202/Domain: serum albumin repeat homology <SA1>  
F;221-394/Domain: serum albumin repeat homology <SA2>  
F;413-592/Domain: serum albumin repeat homology <SA3>





Query Match 78.8%: Score 2446.5: DB 1: Length 607:

Query Match 78.48; Score 2432.5; DB 1; Length 607;



262	DB	VHKECCGDLLECADRADLAKYICENQDPISTYKLECCDPILEKSHCIAKAKDELPA	321
301	QY	DLPSLAADVESKDCKVKNYAEAKDVTGLMFXYEYARHPDYSVLLLRJAKTYETLEKC	360
322	DB	DLNPLEHDEVDKEVCKNKEAKDVTGLTFYBSRRHPDYSVLLLRJAKTYEATLEDC	381
361	QY	CAAADPHCYAKVTFDFPKPLVEEPONILKNCSELPEOLGEYKFFONALVRYTKKVPQVST	420
382	DB	CAKEDPPACVATVDFKQFLVDDEFKNLIKNCSELPEKLEGEYGFONALVRYTKKVPQVST	441
421	QY	PTLVESVNLKGVSKCKKHPEAKNMPCAEDYLSVVLNOLCVLHEKTPVSDRVTCKCTES	480
442	DB	PTLIVEAARKLVGSRCCRRPEERLSCARDYLSLNLARLVCVLHEKTPVSEKVTCKCTES	501
481	QY	LNVNRCPFSALVDENYTVPKFENATFTFHADJCTLTSKEROIKKQFALVELVKEHKPAT	540
502	DB	LNVNRCPFSALTDDYTKPKFVEGCTFTFHADJCTLVPEDEKQIKKQFALVELLKHHPAT	561
541	QY	KEQLKAVMDDFAPVSEKCKCKADKTCFABEGKKLV	576
562	DB	EEQLRTVLGNFAAFVOKKCAAPDHEACFAVEGPKFV	597

ΔO

```

502 LYNRRCPESALTPTYKKEFVEGTFTHADICTLPDEKQIKKQTALVELLKKHPHAT 561
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 KQQLKAVMDPFAAPVEKCKKADDKETCFABEGKKLV 576
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 562 EEQLRTVLGNFAAFVQKCCAAAPDEACFAVEGPKFV 597
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 9  
 JC5838  
 albumin - Mongolian jird  
 C:Species: Meriones unguiculatus (Mongolian jird)  
 C:Date: 05-Mar-1998 #sequence\_revision 13-Mar-1998 #text\_change 19-May-2000  
 C:Accession: JC5838  
 R:Yoshida, K.; Seto-Ohashima, A.; Sinohara, H.  
 DNA Res. 4, 351-354, 1997  
 A:Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis  
 A:Reference number: JC5838; MUID:98116663; PMID:9455485  
 A:Accession: JC5838  
 A:Molecule type: mRNA  
 A:Residues: 1-609 <OS>  
 A:Cross-references: DBJ:AB006197; NID:g2317277; PID:g2317278  
 A:Experimental source: liver  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 F:222-395/Domain: serum albumin repeat homology <SA2>

Db	87	CDKSLHTLFGDKLCSLPTNGEKTADWACCNAKQEPNERCEFLQHKDNDNPOLPFFRAEPD	146
QY	122	VMTAFDHNDEETFLKYLKYIEIARRHPYFYAPELLFTFAKYKAAFTFECQADAAACLLP	181
Db	147	ANCTAFQENASAFNGHYLHEIVARRHPYFTGPELLIYLADTYAVLTCEAADDKGACLLP	206
QY	182	LDELDEBGKASSAKORLCKASLQKGEGRFAKAWARLSSORPKAEFAEYSLKVLDTLTKV	241
Db	207	LDALKEKALVSAVRQLKSSMKKIGERFAKAWARMSQTFFPNADFAEITKLADTLTKV	266
QY	242	HTCECHGDLLECADRADLAKYICENQDSTSSKLTCCEKPLLEKSHCIAEVENDEMPAD	301
Db	267	TOCECHGDLLECADRAELAKYMENQASTSSLQACCDKEMLOKSCTAEVEHDDMPAD	326
QY	302	LPSLAADFVBEKQVCKNFAEKDVLGFMFYEVYARSHPDYSWLLLLRLAKYETVTEKCC	361
Db	327	LPALITADFVEKQVCKNFAEKDVLGTFYLYEVSRHPYSVSLLLRLAKYETVTEKCC	386
QY	362	AAADPHCYAKVTFDEKPLVBEPQNLIKQNCLEFQOLGEBYKFNQNALIVRYTKKQVSTP	421
Db	387	AEADPHACVGHVTFDEKPLVBEPQNLKSNCELYEKLYGFGFQNAVLYRYTKKQVSTP	446

QY 422 TLVEVSRNLGKVGKCKHPEAKRMPCAEDYLSVLNQLCVLHKEKTPVSDRVTRKCTESL 481  
 |||||:||||: || |||||:||||: |||||:||||: |||||:||||: |||||: ||  
 Db 447 TLVEAARSLGRVTHCCALPEKRLPCVEDYLSAILNRVCLLHKEKTPVSEQVTRCCGSL 506  
 QY 482 VNRRPCTSALEVEDTTPVEKFAETFTTHADICTLSEKERQIKKQATVALVELVKHKPKATK 541  
 |||||:||||: |||||:||||: |||||:||||: |||||:||||: |||||: ||  
 Db 507 VERRPCTSALEVEDTTPVEKFAETFTTHADICTLSEKERQIKKQATVALVELVKHKPKATK 566  
 QY 542 EQLKAVMDDAFVEKCKADKCTCFAREGKGLVRAASQAL 583  
 |||||:||||: |||||:||||: |||||:||||: |||||:||||: |||||: ||  
 Db 567 EQLKAVMDDAFVEKCKADKCTCFAREGKGLVRAASQAL 608  
 RESULT 10  
 A05139  
 serum albumin - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 05-Jun-1987 #sequence\_revision 17-Mar-2000 #text\_change 17-Mar-2000  
 C:Accession: A05139; I48638  
 R:Minighetti, P.P.; Law, S.W.; Dugaiczky, A.  
 Mol. Biol. Evol. 2, 347-358, 1985  
 A:Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudogenes  
 A:Reference number: A93055; MUID:88216123; PMID:2452956  
 A:Accession: A05139  
 A:Molecule type: mRNA  
 A:Residues: 1-418 <MIN>  
 A:CROSS-references: GB:M16111; NID:g191764; PIDN:AAA37190.1; PID:g191765  
 R:Boccaccio, C.; Deschattrette, J.; Meunier-Rotival, M.  
 Gene 88, 181-186, 1990  
 A:Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in the genome of the house mouse  
 A:Reference number: I48638; MUID:90269606; PMID:1971802  
 A:Accession: I48638  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 379-453 <BOC>  
 A:CROSS-references: EMBL:X13060; NID:g52939; PIDN:CAA31458.1; PID:g989334  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 C:Keywords: carrier protein; duplication; metal binding; plasma  
 F:1-104/Domain: serum albumin repeat homology (fragment) <SA1>  
 F:123-296/Domain: serum albumin repeat homology <SA2>  
 F:315-453/Domain: serum albumin repeat homology (fragment) <SA3>  
 Query Match 60.0%; Score 1861; DB 2; Length 453;  
 Best Local Similarity 72.2%; Pred. No. 9.5e-116;  
 Matches 327; Conservative 64; Mismatches 62; Indels 0; Gaps 0;  
 QY 75 CTVATLRETVMGADCCAKQEPERNECFLOHKDNNPLNPLVEPEVDVNCFAFHNEETP 134  
 Db 1 CAIPNLRENTGELADCTQEPERNECFLOHKDNNPLNPLVEPEVDVNCFAFHNEETP 60  
 QY 135 LKKYLYETARRHFFYAPELLFAKRYKAFTCCQAADRAACLLPFLDELNDEGKASSA 194  
 Db 61 MGHVLEHVARRRHFFYAPELLFAKRYKAFTCCQAADRAACLLPFLDELNDEGKASSA 120  
 QY 195 KQRLKCSLQKGERAFKAWARLSQFPKAEFAVSKLVYDLIKVHTECHGDLLECA 254  
 Db 121 RQRMKCSMQKGERAFKAWARLSQFPKAEFAVSKLVYDLIKVHTECHGDLLECA 180  
 QY 255 DRRADLAKYICENODSISSEKKECEKPLLEKSHCIAEVENDEMPADLPALAEVFSKD 314  
 Db 181 DRRADLAKYICENODSISSEKKECEKPLLEKSHCIAEVENDEMPADLPALAEVFSKD 240  
 QY 315 VCKNYAKQVFLGMLFYARRHPDYISVLLLLAKTYETTTLEKCCAAADPHCYAKVF 374  
 Db 241 VCKNYAKQVFLGMLFYARRHPDYISVLLLLAKTYETTTLEKCCAAADPHCYAKVF 300  
 QY 375 DEKPLVEEONLKNCELFQOLGEYKFNALLVRYTKKQVOSTPTPLVEVSRNLGKVG 434  
 Db 301 AEFQPLVEEONLKNCELFQOLGEYKFNALLVRYTKKQVOSTPTPLVEVSRNLGKVG 360  
 QY 435 SKCKCHPEAKMCAEDYLSVLNQLCVLHKEKTPVSDRVTRKCTESLVRNRPCTSALEVD 494  
 Db 361 TKCTCLPBDRLPCVEDYLSAILNRVCLLHKEKTPVSEQVTRCCGSLVERRPCTSALEVD 420

QY 495 ETIVPKERNAETFTTHADICTLSEKERQIKKQAT 527  
 |||||:||||: |||||:||||: |||||:||||: |||||:||||: |||||: ||  
 Db 421 ETIVPKERNAETFTTHADICTLSEKERQIKKQAT 453  
 RESULT 11  
 ABCHS  
 serum albumin precursor - chicken  
 C:Species: Gallus gallus (chicken)  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
 C:Accession: S15571; A05078; A13451  
 R:Cassidy, A.I.; Salkild, C.K.; Baverstock, P.; Wallace, J.C.  
 submitted to the EMBL Data Library, July 1991  
 A:Reference number: S15571  
 A:Accession: S15571  
 A:Molecule type: mRNA  
 A:Residues: 1-615 <CRS>  
 A:CROSS-references: EMBL:X60688; NID:g63747; PIDN:CAA43098.1; PID:g63748  
 R:Hache, R.J.G.; Wiskocil, R.; Vasa, M.; Roy, R.N.; Lau, P.C.K.; Deeley, R.G.  
 J. Biol. Chem. 258, 4556-4564, 1983  
 A:Title: The 5' noncoding and flanking regions of the avian very low density apolipoprotein  
 A:Reference number: A05078; MUID:83161037; PMID:6187737  
 A:Accession: A05078  
 A:Molecule type: DNA  
 A:Residues: 1-28 <HAC>  
 A:CROSS-references: GB:V00381; NID:g63038; PIDN:CAA23680.1; PID:g63039  
 R:Rosen, A.M.; Geller, D.M.  
 Biochem. Biophys. Res. Commun. 78, 1060-1066, 1977  
 A:Title: Chicken microsomal albumin: amino terminal sequence of chicken prealbumin.  
 A:Reference number: A13451; MUID:78019943; PMID:911327  
 A:Accession: A13451  
 A:Molecule type: protein  
 A:Residues: 19-23,'M',25-30 <ROS>  
 C:Comment: Serum albumin is synthesized in the liver as prealbumin. It binds copper  
 mones (weak bonds with these hormones promote their transfer across the membranes), t  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 C:Keywords: carrier protein; duplication; metal binding; plasma  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-26/Domain: propeptide #status predicted <PRO>  
 F:27-613/Product: serum albumin #status predicted <MAI>  
 F:32-206/Domain: serum albumin repeat homology <SA1>  
 F:225-398/Domain: serum albumin repeat homology <SA2>  
 F:417-596/Domain: serum albumin repeat homology <SA3>  
 F:30/Binding site: copper (His) #status predicted  
 F:80-89,102-118,117-128,152-197,196-205,228-274,273-281,293-307,306-317,344-389,388-3  
 Query Match 50.2%; Score 1557.5; DB 1; Length 615;  
 Best Local Similarity 46.7%; Pred. No. 1.6e-95;  
 Matches 273; Conservative 118; Mismatches 192; Indels 1; Gaps 1;  
 QY 3 HKSEVAREKDLGRENFKALVLIATAFYLOQCPPEDEHVKLVNEVTEFAKTVADSEANC 62  
 Db 30 HKSEVAREKDLGRENFKALVLIATAFYLOQCPPEDEHVKLVNEVTEFAKTVADSEANC 89  
 QY 63 DKSLETLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPLNPLR-LVPEVD 121  
 Db 90 SKPLPSIILDEICQVEKLRDSTGAMADCCAKQEPERNECFLOHKDNNPLNPLR-LVPEVD 149  
 QY 122 VMCFAFHNEETFLKYLVEIARRHPDYAPELLFAKRYKAFTCCQAADRAACLLP 181  
 Db 150 VMCFAFHNEETFLKYLVEIARRHPDYAPELLFAKRYKAFTCCQAADRAACLLP 209  
 QY 182 LDELNDEGKASSAKORLKASLQKGERAFKAWARLSQFPKAEFAVSKLVYDLIKV 241  
 Db 210 LDELNDEGKASSAKORLKASLQKGERAFKAWARLSQFPKAEFAVSKLVYDLIKV 269  
 QY 242 HTECHGDLLECAADRAADLAKTYICENODSISSEKKECEKPLLEKSHCIAEVENDEMPAD 301  
 Db 270 HTECHGDLLECAADRAADLAKTYICENODSISSEKKECEKPLLEKSHCIAEVENDEMPAD 329  
 QY 302 LPSLAADFVSKDVKCKNYAEAKDVLGMLFYARRHPDYISVLLLLAKTYETTTLEKCC 361  
 |||||:||||: |||||:||||: |||||:||||: |||||:||||: |||||: ||

```

Db 330 LPSIVKYEYIEDKEVCKSPFAGHDFAFYAEFYVYSRRHPFSIQILMIPIANGYESLLEKCC 389
QY 362 AADPHCYAKVDFEKLVEEPONLIKONCELFQOLGEYKFNALLVRYTKVPQVSTP 421
Db 390 KTDNPACVANAQELAQHIKETQDVVTKNCDLLHDHGEADFLASILIRYTKMPQVPTD 449
QY 422 TLVEYSRNLGKVGCKCKHPKAPKMPCAEDYLSVNLQNLVHEKTPVSDRYTKCTESL 481
Db 450 LLETKKMTTIGTKCOOLGDRMACSGEYLSIVHDTCKRQETTPINDNVSCQSQLY 509
QY 482 VNRRPCFSALEYDETYVPEKFNATFTFHADICTLSEKERQIKKOTALVELVHKFKATK 541
Db 510 ANRRPCFTAMGVDTYVPPPNPDMFSEFKLCSAPAEERVGQMKLLINLKRKQMT 569
QY 542 BOLKAVNDFAFAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
Db 570 EQKITADGFTAMVDKCKQSDINTCFGEAGANLVQSRATIGI 613

RESULT 12
JC4258
alpha-fetoprotein precursor - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C:Date: 27-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 20-Aug-1999
C:Accession: JC4258
R:Nishio, H.; Gibbs, P.E.M.; Minghetti, P.P.; Zielinski, R.; Dugaiczky, A.
Gene 162, 213-220, 1995
A:Title: The chimpanzee alpha-fetoprotein-encoding gene shows structural similarity to
A:Reference number: JC4258; MUID:96032345; PMID:7557431
A:Accession: JC4258
A:Molecule type: DNA
A:Residues: 1-609 <NLS>
A:Cross-references: GB:U21916; NID:g841311; PIDN:AA91641.1; PID:g841312
C:Comment: This protein is a plasma protein produced in the fetal and neonatal liver and
o similar properties and structure.
C:Genetics:
A:Gene: afp
A:Map position: 3p
A:Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551/3
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: glycoprotein
F:1-13/Domain: signal sequence #status predicted <SIG>
F:20-609/Product: alpha-fetoprotein #status predicted <NAT>
F:29-202/Domain: serum albumin repeat homology <SAL>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:42,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.4%; Score 1253.5; DB 2; Length 609;
Best Local Similarity 40.1%; Pred. No. 2e-75;
Matches 236; Conservative 115; Mismatches 231; Indels 7; Gaps 3;

QY 3 HKSE-----VAHRFKDLGEENFKALVLIAPQVLOCPEDRVKLVNEVTEFAKTCVAD 57
Db 22 HRNVEYGASILDSYQCTAEINLDTLWIFFAQVQVATYKYSKVMKDALTAIEKPTGDE 81
QY 58 SAENCKSLHTLFGDKLCTVATLRETYGENADCCAKQEPNECPLOHKDNP-NLPLRV 116
Db 82 QSAGCLNQLPAFLEELCKEIELEKYCH-SDCCSSEGRNCHFLAKKFTPASIFPTQ 140
QY 117 RPEYDVMTAFHNDNEETFLKLYEIAHRRHPYFAPLEFFAKRYKKAFTCECCQAADKAA 176
Db 201 CFQZKAATVTKELRESSLNQHACAVMKNEFTGTFQAITVTKLSOKFTKVNFTETOKLVL 260
QY 237 DLTVHVECHGDLLECADRDADLAKYCENQDSISKLKECCERPLEKSHCIAFVEND 296
Db 261 DVAVHHCRCRGLDCLDQDGKMTSYCSQQDTLSNKTCKTECQTLTLTKRQCIIHAEND 320
QY 297 EMPADLPFLAADFVSKDCKNYAEAKDVFGLMFLYVARRHPDYSVLLRLAKTVEIT 356

```

```

Db 321 EPEGLSNLNRFLGDRDFNQFSGGKNFIFLASFVHEYSRRHPOLAVSVILRVAKGYOEL 380
QY 357 LEKCCAAADPHCYAKVDFEKLVEEPONLIKONCELFQOLGEYKFNALLVRYTKVP 416
Db 381 LEKCFOTENPLECODKGEELQYIQESQALAKRSCGLFQKLGLEYLONAEFLVATYTKAP 440
QY 417 QVSTPILVEYSRNLGKVGCKCKHPKAPKMPCAEDYLSVNLQNLVHEKTPVSDRYTKC 476
Db 441 QLTSSSELMAITRKMAATAATCCQLSEDKLLACGGAADIIGHCICIRHETTPVNFVGQCC 500
QY 477 CTESIVNRRPCFSALEYDETYVPEKFNATFTFHADICTLSEKERQIKKOTALVELVHKH 536
Db 501 CTSSVANRRPCFSSLVDETYVPPAFSDDKFIFHDKLQCAQGVALQMKQEFLLNVRKQ 560
QY 537 PKATKQLKAVMDPFAAFVEKCKKADDKETCFABEGKKLVAASQAALGL 585
Db 561 PQITEQLEAVIADFSGLEKCCQGBQVCFABEGQKLSKTRALGV 609

```

## RESULT 13

```

FPHU
alpha-fetoprotein precursor [validated] - human
N:Alternate names: AFP; alpha-1-fetoprotein; alpha-fetoglobulin
C:Species: Homo sapiens (man)
C:Date: 19-Feb-1984 #sequence_revision 19-Feb-1984 #text_change 08-Dec-2000
C:Accession: A26624; S37655; A93961; A91497; A23699; A61480; A90757; A93042;
R:Gibbs, P.E.M.; Zielinski, R.; Boyd, C.; Dugaiczky, A.
Biochemistry 26, 1332-1343, 1987
A:Title: Structure, polymorphism, and novel repeated DNA elements revealed by a comp
A:Reference number: A26624; MUID:87185438; PMID:2436661
A:Accession: A26624
A:Molecule type: DNA
A:Residues: 1-609 <GIB>
A:Cross-references: GB:M16110; NID:g773678; PIDN:AA858754.1; PID:g178236
R:McVey, J.H.; Michaelides, K.; Hansen, L.P.; Ferguson-Smith, M.; Tildeman, S.; Krum
Hum. Mol. Genet. 2, 379-384, 1993
A:Title: A G-A substitution in an HNF I binding site in the human alpha-fetoprotein
A:Reference number: S37655; MUID:93278385; PMID:7684942
A:Accession: S37655
A:Molecule type: DNA
A:Residues: 1-28 <MCV>
A:Cross-references: EMBL:Z19532; NID:g28527; PIDN:CAA79592.1; PID:g28528
A:Note: The authors translated the codon TAA for residue 26 as Thr
R:Morinaga, T.; Sakai, M.; Wegmann, T.G.; Tamaoki, T.
Proc. Natl. Acad. Sci. U.S.A. 80, 4604-4608, 1983
A:Title: Primary structures of human alpha-fetoprotein and its mRNA.
A:Reference number: A93961; MUID:83273664; PMID:6192439
A:Accession: A93961
A:Molecule type: mRNA
A:Residues: 1-609 <MOR>
A:Cross-references: GB:J00077; NID:g311348; PIDN:CAA24758.1; PID:g31351
R:Beattie, W.G.; Dugaiczky, A.
Gene 20, 415-422, 1982
A:Title: Structure and evolution of human alpha-fetoprotein deduced from partial seq
A:Reference number: A91497; MUID:83158778; PMID:6187626
A:Accession: A91497
A:Molecule type: mRNA
A:Residues: 429-556 <BEA>
A:Cross-references: GB:J000076
R:Pucci, P.; Siciliano, R.; Malorni, A.; Marino, G.; Tecce, M.F.; Ceccarini, C.; Ter
Biochemistry 30, 5061-5066, 1991
A:Title: Human alpha-fetoprotein primary structure: a mass spectrometric study.
A:Reference number: A23699; MUID:91242409; PMID:1709810
A:Accession: A23699
A:Molecule type: protein
A:Residues: 19-45; 60-97; 102-107; 122-184; 187-249; 255-489; 507-609 <PUC>
R:Tecce, M.F.; Terrana, B.; Giuliani, M.M.; Ceccarini, C.
J. Nucl. Med. Allied Sci. 34, 213-216, 1990
A:Title: Characterization of in vitro expressed human alpha-fetoprotein as highly re
A:Reference number: A61480; MUID:91225826; PMID:1709209
A:Accession: A61480
A:Molecule type: protein

```

Db 141 VPEVPTSCAYEEDRTFMNFKVYIETIARRHPFIYAPTILLH#AARYDKTIIPSCCKAENAVE 200  
 QY 177 CLLKPLDELDEGKASSAKORLKASLQKTEGRAFAKAWARLSORFPKAEFAEVSKLVT 236  
 Db 201 CFQTKAATVTKELRESSLLNQHACAVMKNFTTFOAITVTKLSOKFTKVNFTIQLVL 260  
 QY 237 DLTQVHTTECHGDLLECCADRADLAKYICENQDSISSKILKECECKPLLEKSHCIAEVEND 296  
 Db 261 DVAHVHCHCRGVDLCLQDGKMSVTC5QDWTLSNKTICECKLTTLERQOCIHAEND 320  
 QY 297 EMPADLPSLAADVESKDYCKNAYAKDVLGFMFLYEARHPDSYVILLRLAKAYEIT 356  
 Db 321 EKPEGLSPNLRNLGRDNPFSGCKENFIASFVHEYSRRHPQLAVSVILLRAKAGQEL 380  
 QY 357 LEKCAAAADPHECYAKVDFEKKPVLVEPQNLIKQNCLEFELQGEYKFFQNALLVYTKKVP 416  
 Db 381 LEKCFQTFENLECODKGEELQKTIQESQALAKRSCGLFKQLGEYLLQNAFLVAYTKAP 440  
 QY 417 QVSTPTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHKTPVSDRVTKC 476  
 Db 441 QLTSESLMATTKMAATAATCCQLSDKLLACGEGAAITIIHGLCIRHEMTPVNPVGVC 500  
 QY 477 CTESLVNRRFCFSALRVDETVYVPKEFNAETFTHADICLTSEKRIQIKKQALVELVKKH 536  
 Db 501 CTSTANRRFCFSLVVDYTVYPAFSDKFIHKDLCQAQGVALTQMKQFLINLVKOR 560  
 QY 537 PRATKEQLKAVMDFFAAFEKCKRADDKTCFAEEGKILVAA5AALGL 585  
 Db 561 PQITEQLEAVIADFSLLLEKCCQGBQEVCFEAEGOKLISKTRALGV 609  
 RESULT 14  
 FPGO  
 alpha-fetoprotein precursor - gorilla  
 C:Species: Gorilla gorilla (gorilla)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
 C:Accession: A37970  
 R:Ryan, S.C.; Zielinski, R.; Dugaiczky, A.  
 Genomics 9, 60-72, 1991  
 A:Title: Structure of the gorilla alpha-fetoprotein gene and the divergence of  
 A:Reference number: A37970; MUID:91169517; PMID:1706310  
 A:Accession: A37970  
 A:Molecule type: DNA  
 A:Residues: 1-609 <R>A>  
 A:Cross-references: GB:M38272; NID:g817963; PIDN:AAA73520.1; PID:g177041  
 C:Genetics:  
 A:Map position: 4q11-12  
 A:Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2;  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 K:Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-609/Product: alpha-fetoprotein #status predicted <MANT>  
 F:29-202/Domain: serum albumin repeat homology <SAL>  
 F:221-394/Domain: serum albumin repeat homology <SAR2>  
 F:413-592/Domain: serum albumin repeat homology <SAR3>  
 F:22/Binding site: copper (His) #status predicted  
 F:99-114,113-124,148-193,192-270,269-277,289-303,302-313,384-393,416-417  
 F:249/Binding site: bilirubin (Lys) #status predicted  
 F:251/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 40.0%; Score 1242.5; DB 1; Length 609;  
 Best Local Similarity 39.6%; Pred. No. 1.1e-74;  
 Matches 233; Conservative 117; Mismatches 232; Indels 7; Gaps 3  
 QY 3 HKSE-----VAHFKDLDGEENFKALVLIATRAQVILQOCPEFDHVKLVNEVTEFAKCVADE 57  
 Db 22 HRNEYGTASTLDSTQCTAEISLADIATITFAQFQAEATKYKSMVKDAITAEKPTGDE 81  
 QY 58 SAENCKDSLHTLGDKLCVATLRETYGEMADCCAKOEPERNECFLOHKDNP--NLPRLV 116  
 Db 82 OSAGLENQLPAFLELCEKEILEKYG--ISDCSSSEGRNCFIAHKFTPASIFLQ 140  
 QY 117 RPEVDNMCTAFHDNEETFLKKLYLRIARHPFIYAPPELLFFFAKRYKAATFCCQAADKAA 176



Search completed: August 31, 2003, 16:34:37  
Job time : 45 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2003, 16:33:10 ; Search time 25 Seconds  
(without alignments)  
1100.425 Million cell updates/sec

Title: US-09-833-041-18

Perfect score: 3103

Sequence: 1 DAKSEVAHFQDLGEENFK.....TCFABEGKKLVASQAALGL 585

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3103	100.0	609	1	ALBU_HUMAN
2	2942	94.8	600	1	ALBU_MACMU
3	2620	84.4	608	1	ALBU_FELCA
4	2562	82.6	608	1	ALBU_CANFA
5	2475.5	79.8	607	1	ALBU_HORSE
6	2450.5	79.0	607	1	ALBU_BOVIN
7	2446	78.8	608	1	ALBU_RABIT
8	2432.5	78.4	607	1	ALBU_SHEEP
9	2426	78.2	608	1	ALBU_RAT
10	2411.5	77.7	605	1	ALBU_PIG
11	2387	76.9	609	1	ALBU_MERON
12	2378	76.6	608	1	ALBU_MOUSE
13	1557.5	50.2	615	1	ALBU_CHICK
14	1253.5	40.4	609	1	FETA_PANTR
15	1249.5	40.3	609	1	FETA_HUMAN
16	1242.5	40.0	609	1	FETA_GORGO
17	1205	38.8	607	1	ALB2_XENLA
18	1200	38.7	609	1	FETA_HORSE
19	1164.5	37.5	606	1	ALB1_XENLA
20	1084	34.9	605	1	FETA_MOUSE
21	1067	34.4	611	1	FETA_RAT
22	1055	34.0	599	1	AFAM_HUMAN
23	944	30.4	611	1	AFAM_MOUSE
24	928	29.9	608	1	AFAM_RAT
25	747.5	24.1	608	1	ALB1_SALSA
26	742.5	23.9	608	1	ALB2_SALSA
27	699	22.5	382	1	ALBU_RANCA
28	440.5	14.2	1423	1	ALBU_PETMA
29	386	12.4	474	1	VTDB_HUMAN
30	381	12.3	476	1	VTDB_RAT
31	378	12.2	476	1	VTDB_RABIT
32	372	12.0	472	1	VTDB_MOUSE
33	151.5	4.9	1605	1	RRE1_MOUSE

## RESULT 1

ALBU_HUMAN	1	SNEL_HUMAN	8797	4.7	144.5	34
AC	P02768; O95574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9UJZ0;	1	1410	4.5	138.5	35
DT	21-JUL-1986 (Rel. 01, Created)	1	1391	4.3	133.5	36
DT	01-APR-1990 (Rel. 14, Last sequence update)	1	2230	4.3	132.5	37
DT	15-SEP-2003 (Rel. 42, Last annotation update)	1	1972	4.2	129.5	38
DE	Serum albumin precursor.	1	3210	4.2	129	39
GN	ALB.	1	1005	4.1	128	40
OS	Homo sapiens (Human).	1	1972	4.1	126.5	41
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	1	1189	4.1	126	42
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	1	3259	4.1	126	43
OX	NCBI_TaxID=9606;	1	3038	4.0	125	44
RN	[1]	1	1790	4.0	124.5	45
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=86196112; PubMed=3009475;					
RA	Minghetti P.P., Rufner D.E., Kuang W.J., Dennison O.E., Hawkins J.W., Beattie W.G., Dugaiczky A.;					
RT	"Molecular structure of the human albumin gene is revealed by nucleotide sequence within g11-22 of chromosome 4.";					
RL	J. Biol. Chem. 261:6747-6757(1986).					
RN	[2]					
RP	SEQUENCE FROM N.A., AND VARIANT LYS-420.					
RX	MEDLINE=82081882; PubMed=6171778;					
RA	Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M., Najarian R.C., Seeburg P.H., Wion K.L.;					
RT	"The sequence of human serum albumin cDNA and its expression in E. coli.";					
RL	Nucleic Acids Res. 9:6103-6114(1981).					
RN	[3]					
RP	SEQUENCE FROM N.A., AND VARIANT GLY-121.					
RX	MEDLINE=82105994; PubMed=6275391;					
RA	Dugaiczky A., Law S.W., Dennison O.E.;					
RT	"Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.";					
RL	Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).					
RN	[4]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=Liver;					
RA	Xang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;					
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.					
RN	[5]					
RP	SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).					
RC	TISSUE=Fetal liver;					
RA	Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y., Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;					
RT	"Functional prediction of the coding sequences of 121 new genes deduced by analysis of cDNA clones from human fetal liver.";					
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.					
RN	[6]					
RP	SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.					
RA	Huang M.C., Wu H.T.;					
RT	"The cDNA sequences of human serum albumin.";					
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.					
RN	[7]					
RP	SEQUENCE FROM N.A.					

Q8af91 homo sapien  
Q9p2e9 homo sapien  
Q08696 drosophila  
Q13439 homo sapien  
P35749 homo sapien  
P49454 homo sapien  
Q58718 methanococ  
P35748 oryctolagus  
Q90988 gallus gall  
Q14789 homo sapien  
O75962 homo sapien  
P25386 saccharomyc



RC TISSUE=Liver, and Skeletal muscle;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
RA Atschul S.F., Zeeberg B., Suecow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RA "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [8]  
RP SEQUENCE OF 25-609.  
RX MEDLINE=76187907; PubMed=1225573;  
RA Meloun B., Moravsek L., Kostka V.;  
RT "Complete amino acid sequence of human serum albumin.";  
RL FEBS Lett. 58:134-137(1975).  
RN [9]  
RP SEQUENCE OF 25-609.  
RA Brown J.R., Shockley P., Behrens P.Q.;  
RL (In) Bing D.H. (eds.);  
RL The chemistry and physiology of the human plasma proteins, pp.23-40,  
RL Pergamon Press, New York (1979).  
RN [10]  
RP SEQUENCE OF 1-455 FROM N.A.  
RC TISSUE=Liver;  
RA Menaya J., Parrilla R., Ayuso M.S.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
RN [11]  
RP SEQUENCE OF 1-26 FROM N.A.  
RX MEDLINE=86140099; PubMed=2419329;  
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;  
RT "The human albumin gene. Characterization of the 5' and 3' flanking  
RL regions and the polymorphic gene transcripts.";  
RL J. Biol. Chem. 261:3244-3251(1986).  
RN [12]  
RP SEQUENCE OF 222-229.  
RX MEDLINE=76257808; PubMed=955075;  
RA Walker J.E.;  
RT "Lysine residue 199 of human serum albumin is modified by  
RL acetylsalicylic acid.";  
RL FEBS Lett. 66:173-175(1976).  
RN [13]  
RP SEQUENCE OF 25-44 AND 480-499.  
RC TISSUE=Heart;  
RX MEDLINE=95203287; PubMed=7895732;  
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
RT "The human myocardial two-dimensional gel protein database: update  
RL 1994.";  
RL Electrophoresis 15:1459-1465(1994).  
RN [14]  
RP DISULFIDE BONDS.  
RA Sabar M.A., Stockbauer P., Moravsek L., Meloun B.;  
RT "Disulfide bonds in human serum albumin.";  
RL Collect. Czech. Chem. Commun. 42:564-579(1977).  
RN [15]  
RP BILIRUBIN-BINDING SITE.  
RX MEDLINE=78186630; PubMed=656055;  
RA Jacobsen C.;  
RT "Lysine residue 240 of human serum albumin is involved in high-  
RL affinity binding of bilirubin.";  
RL Biochem. J. 171:453-459(1978).

[16]  
RN VARIANT CANTERBURY ASN-337.  
RX MEDLINE=87157744; PubMed=3828358;  
RA Brennan S.O., Herbert P.;  
RT "Albumin Canterbury (313 lys->asn). A point mutation in the second  
RL domain of serum albumin.";  
RL Biochim. Biophys. Acta 912:191-197(1987).  
RN [17]  
RP VARIANTS NAG-2 AND NAG-3.  
RX MEDLINE=88068523; PubMed=3479777;  
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,  
RA Satoh C., Neel J.V.;  
RT "Amino acid substitutions in inherited albumin variants from  
RL Amerindian and Japanese populations.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).  
RN [18]  
RP VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.  
RX MEDLINE=89345611; PubMed=2762316;  
RA Arai K., Madison J., Huss K., Ishioka N., Satoh C., Fujita M.,  
RA Neel J.V., Sakurabayashi I., Putnam F.W.;  
RT "Point substitutions in Japanese alloalbumins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).  
RN [19]  
RP VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.  
RX MEDLINE=90115905; PubMed=2404284;  
RA Arai K., Madison J., Shimizu A., Putnam F.W.;  
RT "Point substitutions in albumin genetic variants from Asia.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).  
RN [20]  
RP DESCRIPTION OF VARIANT REDHILL.  
RX MEDLINE=90115852; PubMed=2104980;  
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;  
RT "Albumin Redhill (-1 Arg, 320 Ala->Thr): a glycoprotein variant of  
RL human serum albumin whose precursor has an aberrant signal peptidase  
RL cleavage site.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).  
RN [21]  
RP VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.  
RX MEDLINE=91062352; PubMed=2247440;  
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,  
RA Watkins S., Putnam F.W.;  
RT "Mutations in genetic variants of human serum albumin found in  
RL Italy.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).  
RN [22]  
RP VARIANT VENEZIA.  
RX MEDLINE=91296740; PubMed=2068071;  
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,  
RA Minchiotti L., Putnam F.W.;  
RT "A donor splice mutation and a single-base deletion produce two  
RL carboxyl-terminal variants of human serum albumin.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).  
RN [23]  
RP VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;  
RX KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.  
RX MEDLINE=92052189; PubMed=1946412;  
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,  
RA Matsuda Y.-I., Amaki I., Putnam F.W.;  
RT "Genetic variants of serum albumin in Americans and Japanese.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).  
RN [24]  
RP VARIANT CASEBROOK ASN-518.  
RX MEDLINE=91316157; PubMed=1859851;  
RA Peach R.J., Brennan S.O.;  
RT "Structural characterization of a glycoprotein variant of human serum  
RL albumin: albumin Casebrook (494 Asp->Asn).";  
RL Biochim. Biophys. Acta 1097:49-54(1991).  
RN [25]  
RP VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.  
RX MEDLINE=92190239; PubMed=1347703;  
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,  
RA Rochu D., Porta F.;  
RT "Two alloalbumins with identical electrophoretic mobility are produced

Query Match 100.0%; Score 3103; DB 1; Length 609;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-194;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFDLGEENFKALVLIAPAYLOQCPEDHVKLVNEVTEFAKTCVADSEAE 60  
 DB 25 DAHSEVAHRFDLGEENFKALVLIAPAYLOQCPEDHVKLVNEVTEFAKTCVADSEAE 84  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPRV 120  
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPRV 144  
 QY 121 DVMTAFPHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAFTCCQAADKAACLLP 180  
 DB 145 DVMTAFPHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAFTCCQAADKAACLLP 204  
 QY 181 KLDELDEGKASSAQRLKASLOKFGERAFAKAWARLSQFFKAEFAVSKLVTDLJK 240  
 DB 205 KLDELDEGKASSAQRLKASLOKFGERAFAKAWARLSQFFKAEFAVSKLVTDLJK 264  
 QY 241 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEYENDEMPA 300  
 DB 265 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEYENDEMPA 324  
 QY 301 DLPSLAADFVSKDCKNYAKADVFLGMFLYIYARRHPDYSVILLRLAKTETTLK 360  
 DB 325 DLPSLAADFVSKDCKNYAKADVFLGMFLYIYARRHPDYSVILLRLAKTETTLK 384  
 QY 361 CAADAPHECYAKVDFEFLVEEPQNLKONCELFQELGEYKFNALLVRYTKVPQVST 420  
 DB 385 CAADAPHECYAKVDFEFLVEEPQNLKONCELFQELGEYKFNALLVRYTKVPQVST 444  
 QY 421 PTLVSVSRNLKGVSKCKHPEAKMPCAEADYLSVNLQCVLHEKTPVSDRYTKCCTES 480  
 DB 445 PTLVSVSRNLKGVSKCKHPEAKMPCAEADYLSVNLQCVLHEKTPVSDRYTKCCTES 504  
 QY 481 LVNRRCFSALEVDYETVYKFEFNAETTFHADICTLSEKEROIKQALVELYKHKPKAT 540  
 DB 505 LVNRRCFSALEVDYETVYKFEFNAETTFHADICTLSEKEROIKQALVELYKHKPKAT 564  
 QY 541 KEQLKAVMDFFAFAVEKCKADKTCFAEKGKLVAAASQAALGL 585  
 DB 565 KEQLKAVMDFFAFAVEKCKADKTCFAEKGKLVAAASQAALGL 609

RESULT 2  
 ID ALBU\_MACMU STANDARD; PRG; 600 AA.  
 AC Q28522;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Serum albumin precursor (Fragment).  
 GN ALB.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93211971; PubMed=8460152;  
 RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,  
 RA Dwulet J., Putnam F.W.;  
 RT "cDNA and protein sequence of polymorphic macaque albumins that differ  
 in bilirubin binding.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 hormones, bilirubin and drugs. Its main function is the regulation  
 of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation.  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: M90463; AAA36906.1; --  
 DR PIR: A47391; A47391.  
 DR HSSP: P02768; 1E7B.  
 DR InterPro: IPR000264; Serum\_albumin.  
 DR Pfam: PF00273; Transport\_prot; 3.  
 DR PRINTS: PR00802; SERUMALBUMIN.  
 DR PRODOM: PD002486; Serum\_albumin; 1.  
 DR SMART: SM00103; ALBUMIN; 3.  
 DR PROSITE: PS00212; ALBUMIN; 3.  
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
 FT METAL 1  
 FT SIGNAL <1 10 BY SIMILARITY.  
 FT PROPEP 11 16 BY SIMILARITY.  
 FT CHAIN 17 600 SERUM ALBUMIN.  
 FT DOMAIN 17 197 ALBUMIN 1.  
 FT DOMAIN 204 389 ALBUMIN 2.  
 FT DOMAIN 396 587 ALBUMIN 3.  
 FT METAL 19 19 COPPER (BY SIMILARITY).  
 FT BINDING 256 256 BILIRUBIN (POTENTIAL).  
 FT DISULFID 69 78 BY SIMILARITY.  
 FT DISULFID 91 107 BY SIMILARITY.  
 FT DISULFID 106 117 BY SIMILARITY.  
 FT DISULFID 140 185 BY SIMILARITY.  
 FT DISULFID 184 193 BY SIMILARITY.  
 FT DISULFID 216 262 BY SIMILARITY.  
 FT DISULFID 261 269 BY SIMILARITY.  
 FT DISULFID 281 295 BY SIMILARITY.  
 FT DISULFID 294 305 BY SIMILARITY.  
 FT DISULFID 332 377 BY SIMILARITY.  
 FT DISULFID 376 385 BY SIMILARITY.  
 FT DISULFID 408 454 BY SIMILARITY.  
 FT DISULFID 453 464 BY SIMILARITY.  
 FT DISULFID 477 493 BY SIMILARITY.  
 FT DISULFID 492 503 BY SIMILARITY.  
 FT DISULFID 530 575 BY SIMILARITY.  
 FT DISULFID 574 583 BY SIMILARITY.  
 SQ SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;  
 Query Match 94.8%; Score 2942; DB 1; Length 600;  
 Best Local Similarity 93.5%; Pred. No. 4.8e-184;  
 Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFDLGEENFKALVLIAPAYLOQCPEDHVKLVNEVTEFAKTCVADSEAE 60  
 DB 17 DTHKSEVAHRFDLGEENFKALVLIAPAYLOQCPEDHVKLVNEVTEFAKTCVADSEAE 76  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPRV 120  
 DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPRV 136  
 QY 121 DVMTAFPHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAFTCCQAADKAACLLP 180  
 DB 137 DVMTAFPHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAFTCCQAADKAACLLP 196  
 QY 181 KLDELDEGKASSAQRLKASLOKFGERAFAKAWARLSQFFKAEFAVSKLVTDLJK 240  
 DB 197 KLDELDEGKASSAQRLKASLOKFGERAFAKAWARLSQFFKAEFAVSKLVTDLJK 256  
 QY 241 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEYENDEMPA 300  
 DB 257 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEYENDEMPA 316



DE GN Serum albumin precursor (Allergen Can f 3).

OS ALB.

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OX NCBI\_TaxID=9615;

RN [1]

RN SEQUENCE FROM N.A.

RP STRAIN=Beagle; TISSUE=Liver;

RA Hilger C.;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RN SEQUENCE FROM N.A.

RP TISSUE=Liver;

RC MEDLINE=20148667; PubMed=10669848;

RA Pandjaitan B., Swoboda I., Brandesky-Pichler F., Rumpold H.,

RA Valenta R., Spitzauer S.;

RT "Escherichia coli expression and purification of recombinant dog

RT albumin, a cross-reactive animal allergen.";

RL J. Allergy Clin. Immunol. 105:279-285(2000).

RN [3]

RN SEQUENCE OF 25-48.

RP MEDLINE=75011422; PubMed=4414013;

RA Dixon J.W., Sarkar B.;

RT "Isolation, amino acid sequence and copper(II)-binding properties of

RT peptide (1-24) of dog serum albumin.";

RL J. Biol. Chem. 249:5872-5877(1974).

RN [4]

RN SEQUENCE OF 25-38.

RP TISSUE=Heart;

RC MEDLINE=98163340; PubMed=9504812;

RA Dunn M.J., Corbett J.M., Wheeler C.H.;

RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of

RT dog heart proteins.";

RL Electrophoresis 18:2795-2802(1997).

RN [5]

RN SEQUENCE OF 215-478 FROM N.A.

RC TISSUE=Salivary gland;

RA MEDLINE=94201492; PubMed=7512102;

RA Spitzauer S., Schweizer C., Sperr W.R., Pandjaitan B., Valent P.,

RA Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;

RT "Molecular characterization of dog albumin as a cross-reactive

RT allergen.";

RL J. Allergy Clin. Immunol. 93:614-627(1994).

CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good

CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,

CC hormones, bilirubin and drugs. Its main function is the regulation

CC of the colloidal osmotic pressure of blood.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.

CC -1- SIMILARITY: Contains 3 albumin domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AJ133489; CAB64867.1; -

DR EMBL; Y17737; CA76841.1; -

DR EMBL; S72946; AAB30434.1; -

DR HSP; P02768; 1E7B.

DR HSC-2DPAGE; P49822; DOG.

DR InterPro: IPR000264; Serum\_albumin.

DR Pfam: PF00273; transprot\_prot; 3.

DR PRINTS; PR00802; SERUMALBUMIN.

DR PRODOM; PD002486; Serum\_albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 3.

DR Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.

KW

FT SIGNAL 1 18

FT PROPEP 19 24

FT CHAIN 25 608

FT DOMAIN 25 205

FT DOMAIN 212 397

FT DOMAIN 404 595

FT METAL 27 27

FT DISULFID 77 86

FT DISULFID 99 115

FT DISULFID 114 125

FT DISULFID 148 193

FT DISULFID 192 201

FT DISULFID 224 270

FT DISULFID 269 277

FT DISULFID 289 303

FT DISULFID 302 313

FT DISULFID 340 385

FT DISULFID 384 393

FT DISULFID 416 462

FT DISULFID 461 472

FT DISULFID 485 501

FT DISULFID 500 511

FT DISULFID 538 583

FT DISULFID 582 591

FT CONFLICT 1 26

FT CONFLICT 146 146

FT CONFLICT 206 206

FT CONFLICT 349 349

FT CONFLICT 359 359

FT CONFLICT 448 448

FT CONFLICT 474 474

SQ SEQUENCE 608 AA; 68606 MW; 3CFIC8F7DD8FC06 CRC64;

Query Match 82.6%; Score 2562; DB 1; Length 608;

Best Local Similarity 79.8%; Pred. No. 2.4e-159;

Matches 465; Conservative 57; Mismatches 61; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGLGEENFKALVLIAPAOYLQCCPEDEHVKLVNEYTEFAKTCVADESAB 60

DB 25 EAYKSEIAHRYNDLGEHFRGLVLFVAFSOLQCCPEDEHVKLVNEYTEFAKTCVADESAB 84

QY 61 NCKSLHTLFGDKLCTVATLETYGENADCCACQEPERNECFLOHKDDNPNLPLVRPVP 120

DB 85 NCKSLHTLFGDKLCTVATLETYGENADCCACQEPERNECFLOHKDDNPNLPLVRPVP 144

QY 121 DVMTAFDHNETFLKKLYLIIARRHPYFYAPPELLFFAKRYKAAFTCCQAAKACILP 180

DB 145 DALCAAPQDNEQLFLGKLYLIIARRHPYFYAPPELLFYAQYKGVFAECQAAKACILP 204

QY 181 KIDELRDEGKASSAKORLKCSAQKGFGERAFKAWARLSQRFKAEFAVSKLVDLTK 240

DB 205 KIEALREKVLSSAKERFKCSAQKGFGERAFKAWARLSQRFKAEFAVSKLVDLTK 264

QY 241 VHTCCGHDLLCADDADLAKYTCENQDSISSKKECCERPLEKSHCHTAFTVNDMPA 300

DB 265 VHKCCGHDLLCADDADLAKYTCENQDSISSKKECCERPLEKSHCHTAFTVNDMPA 324

QY 301 DLPSLAADFVESKCKVNAKADKVFGLGMFLYEXARRHPDYSVLLLRKLTAKTETLEKC 360

DB 325 DLPSLAADFVESKCKVNAKADKVFGLGMFLYEXARRHPDYSVLLLRKLTAKTETLEKC 384

QY 361 CAADPHECYAKVFDEFKPLVEEPQNLIKQNCLEFQELGEYKFQNALLVYTKKPVQST 420

DB 385 CATDDPTCYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFQNALLVYTKKPVQST 444

QY 421 PTLVSVRNLGKVGSKCKHPKAPKPCADYLSVVLNQLVLHEKTPVSDRVTKCCTES 480

DB 445 PTLVSVRNLGKVGSKCKHPKAPKPCADYLSVVLNQLVLHEKTPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALVEDETYYPKEFNATETTFHADICTLSEKQIKKOTALVELVKKHPAT 540

DB 505 LVNRRPCFSALVEDETYYPKEFNATETTFHADICTLSEKQIKKOTALVELVKKHPAT 564

QY 541 KEOLKAVMDFAAFVEKCKADKCTCFABEGKLVAAQAAL 583  
 D 565 DBOLKTVMGDFGAFVEKCAENKGCSEBGPKLVAQAAL 607

## RESULT 5

ALBU\_HORSE STANDARD; PRT; 607 AA.  
 AC P35747;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Serum albumin precursor.  
 GN ALB.  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).  
 RC TISSUE=Liver;  
 RX MEDLINE=93345495; PubMed=8344282;  
 RA HO J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;  
 RT "X-ray and primary structure of horse serum albumin (Equus caballus)  
 at 0.27-nm resolution.";  
 RL Eur. J. Biochem. 215:205-212(1993).  
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 hormones, bilirubin and drugs. Its main function is the regulation  
 of the colloidal osmotic pressure of blood.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Plasma.  
 CC -1- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.  
 CC -1- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; X74045; CAA52194.1; -  
 DR F1R; S34053; ABHOS.  
 DR HSP; P02768; 1E7B.  
 DR InterPro: IPR000264; Serum\_albumin.  
 DR Pfam: PF00273; transport\_prot; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum\_albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
 FT SIGNAL 1 18 BY SIMILARITY.  
 FT PROPEP 19 24 BY SIMILARITY.  
 FT CHAIN 25 607 SERUM ALBUMIN.  
 FT DOMAIN 25 204 ALBUMIN 1.  
 FT DOMAIN 211 396 ALBUMIN 2.  
 FT DOMAIN 403 594 ALBUMIN 3.  
 FT METAL 27 27 COPPER (BY SIMILARITY).  
 FT DISULFID 77 86  
 FT DISULFID 99 115  
 FT DISULFID 114 125  
 FT DISULFID 147 192  
 FT DISULFID 191 200  
 FT DISULFID 223 269  
 FT DISULFID 268 276  
 FT DISULFID 288 302  
 FT DISULFID 301 312  
 FT DISULFID 339 384  
 FT DISULFID 383 392  
 FT DISULFID 415 461

FT DISULFID 460 471  
 FT DISULFID 484 500  
 FT DISULFID 499 510  
 FT DISULFID 537 582  
 FT DISULFID 581 590  
 SQ SEQUENCE 607 AA; 68598 MW; 256F6E830A1B90C5 CRC64;  
 Query Match 79.8%; Score 2475.5; DB 1; Length 607;  
 Best Local Similarity 76.3%; Pred. No. 9.8e-154;  
 Matches 445; Conservative 69; Mismatches 68; Indels 1; Gaps 1;  
 QY 1 DAKSEVAHREKDLGEENFKALVLIAPAOYLOQCPPEDEHVKLVEVTEFAKTCVADESAE 60  
 D 25 DTHKSEIAHRENDLGEKHFGLVLFVAFSQLOQCPPEDEHVKLVEVTEFAKCAADESAE 84  
 QY 61 NDKSLHTLFGDKLCTVATRLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120  
 D 85 NDKSLHTLFGDKLCTVATRLRATYGEADCCQEPERNECFLOHKDDNPNLPKL-KPEP 143  
 QY 121 DVMCTAFHDNEETFLKYLIELARHDFYFAPELLFFAKRYKAAFTCCOAAADKACILP 180  
 D 144 DAQCAAFQEDDPDKFLGKLYEVARRHDFYFYPELLPHEAEYKADFTCECPADKLAELP 203  
 QY 181 KLDRLDEGKASSAKQRLKCSLOKFGERAFKAWAYARLSQRFPAEFAEYKIVTDLTK 240  
 D 204 KLDALKERILLSSAKERLKCSSFQNGERAVKAWARSQRFPAEFAEYKIVTDLTK 263  
 QY 241 VHTCCGGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCAEYENDEMA 300  
 D 264 VHKCCGGDLLECADRADLAKYICEHODSISGLKACODKPLLOKSHCAEYKEDDPLS 323  
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFELYETARRHEDYSVLLLLAKARTYETTERC 360  
 D 324 DIPALAADFAEDKEICKYKDAKDFGLTFLYETARRHEDYSVLLLLAKARTYETTERC 383  
 QY 361 CAADPHECVAKVDFEPLVEEPONLKNCEIFPOLGEYFQNALIVRTTKYQVST 420  
 D 384 CAADPHECVAKVDFEPLVEEPONLKNCEIFPOLGEYFQNALIVRTTKYQVST 443  
 QY 421 PTLVEVERNLKGVSKCKKHPKAPKMPKADYLSVNLNOLCVLHKEKTEVSDRVTKCTES 480  
 D 444 PTLVEVERNLKGVSKCKKHPKAPKMPKADYLSVNLNOLCVLHKEKTEVSDRVTKCTES 503  
 QY 481 LVNRRPCFSALEVDETYVPEFNAETFTTHADICTISEKEROIKKQATLVALKHKKPKAT 540  
 D 504 LAERRPCFSALEVDETYVPEFNAETFTTHADICTISEKEROIKKQATLVALKHKKPKAT 563  
 QY 541 KEOLKAVMDFAAFVEKCKADKCTCFABEGKLVAAQAAL 583  
 D 564 DBOLKTVMGDFGAFVEKCAENKGCSEBGPKLVAQAAL 606

## RESULT 6

ALBU\_BOVIN STANDARD; PRT; 607 AA.  
 AC P02769; O02787;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Serum albumin precursor (Allergen Bos d 6).  
 GN ALB.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Holowachuk E.W., Stoltenborg J.K., Reed R.G., Peters T. Jr.;  
 RL Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANT THR-214.  
 RC TISSUE=Liver;

[illegible]

FT	CHAIN	25	608	SERUM ALBUMIN.
FT	DOMAIN	25	205	ALBUMIN 1.
FT	DOMAIN	212	397	ALBUMIN 2.
FT	DOMAIN	404	595	ALBUMIN 3.
FT	METAL	27	27	COPPER.
FT	DISULFID	77	86	BY SIMILARITY.
FT	DISULFID	99	115	BY SIMILARITY.
FT	DISULFID	114	125	BY SIMILARITY.
FT	DISULFID	148	193	BY SIMILARITY.
FT	DISULFID	192	201	BY SIMILARITY.
FT	DISULFID	224	270	BY SIMILARITY.
FT	DISULFID	269	277	BY SIMILARITY.
FT	DISULFID	289	303	BY SIMILARITY.
FT	DISULFID	302	313	BY SIMILARITY.
FT	DISULFID	340	385	BY SIMILARITY.
FT	DISULFID	384	393	BY SIMILARITY.
FT	DISULFID	416	462	BY SIMILARITY.
FT	DISULFID	461	472	BY SIMILARITY.
FT	DISULFID	485	501	BY SIMILARITY.
FT	DISULFID	500	511	BY SIMILARITY.
FT	DISULFID	538	583	BY SIMILARITY.
FT	DISULFID	582	591	BY SIMILARITY.
SO	SEQUENCE	608	68914 MW: CF5E92647AAFE9A2 CRC64:	
		AA:	591	

0; Gaps 0;

275

TCVADESAE 60

1111

ACVADESAA 84.

T BPT WPDW 120

027 ATJATJAT

LPPEARPEA 144

ADKACCLLP 180

[illegible]

ADKGACTP 204

CSKT.VTDT.TK 240

[illegible]

SKITVTDLTK 264

EVENDEMPA 300

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

GLHNDEDTA 324

TYPESET BY K. S. 360

1111

AYEATLKC 384

TKKVPQVST 420

EDUCATION

И.С.А.Н.П.А.В.У.Т.

RVTKCCTES 480

9	
8	
7	
6	—
5	:
4	—
3	—
2	—
1	:

KVTKCCSES 504

LVKHKPKAT 540

**Figure 1**

LVKHKPHAT 564







FT	DISULFID	461	BY SIMILARITY.
FT	DISULFID	485	BY SIMILARITY.
FT	DISULFID	501	BY SIMILARITY.
FT	DISULFID	511	BY SIMILARITY.
FT	DISULFID	538	BY SIMILARITY.
FT	DISULFID	583	BY SIMILARITY.
FT	DISULFID	582	BY SIMILARITY.
FT	VARIANT	262	V -> L
FT	CONFLICT	174	Y -> L (IN REF. 5).
FT	SEQUENCE	174	
SQL	SEQUENCE	608 AA; 68718 MW; 5BB497A282411AB7 CRC64;	
Query Match			
Best Local Similarity 73.4%; Score 2426; DB 1; Length 608;			
Matches 428; Conservative 82; Mismatches 73; Indels 0; Gaps 0;			
QY	1	DAKSEVAFHFDLGERNFKALVILAFQAQYLOCCFEDHVLKLVNEVTFAKTCAVDSAE	60
DB	25	EAKSETAHRFDLGEQFKGLVILAFSQYLOCPYEEHILVQEVDFFAKTCAVDNAE	84
QY	61	NDCKSLHTLFGDKLCTVATLRETYGEMADCCAKQKPERNECFLOHKDDNPNLPRIVR	120
DB	85	NDCKSLHTLFGDKLCAIPKLRDNYGELADCCAKQKPERNECFLOHKDDNPNLPRFORPE	144
QY	121	DVMCTAPHONETFKYLVIETARRHPYFYAPPELLFAKRYKAAFTCCQADKAACLLP	180
DB	145	EAMCTFQENPTSTFLGHVLSARRHPYFYAPPELLFYAEKTYNEVLTCQCTESKAACLP	204
QY	181	KLDRLRDEGKASAKORLKASLOKGERAFKAWAVARLSORFFKAFVSKLVDTLTK	240
DB	205	KLDVKEKALVAVRQRMKCSMORGERAFKAWAVARMSORFFNAFAEITKLAIDVTK	264
QY	241	VHTCECHGLLECADRADLAKYICENQDSISKLEKCEKPLLEKSHCTIAEVNDEMPA	300
DB	265	INKECCHGLLECADRAELAKYMCENQATISKLAQCDKFPVLOKSOCLAEHNDIPA	324
QY	301	DLPSLAADFVESKDYCKNVAEAKDVTFGLMFLEYARRHPDYSVVLLRLAKTYETLEKC	360
DB	325	DLPSLAADFVEDKCYKNVAEAKDVTFGLTFLEYARRHPDYSVLSLLRLAKTYEATLEKC	384
QY	361	CAAADPHECYAKVDFEFLPVEEPQNLIKONCELEFQELGEYKFONALLVRYTKKPVOST	420
DB	385	CAEGDPACYGTVLAEFQPLVEEPKLVKNCLEYKLEGEYGFQNAVLYRTQKAPVST	444
QY	421	PTLVEVSRNLKGVSKCKHPDEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRTVKCTES	480
DB	445	PTLVEARNLGRVGTKCTLPQARLPCVDEYLSALINLCLVHLHEKTPVSEKVTKCCSG	504
QY	481	LVNRRPCFSALVEYDEYTPVKFNAETTFHADICTLSEKEROIKKQALVELVHKRPKAT	540
DB	505	LVRRPCFSALVDEYTPVKFNAETTFHSDICTLPDKERQIKKQALVLAELVHKRPKAT	564
QY	541	KEQLKAVMDFFRAVFEKCKCKADDTCTCFADGEKKLVASQAAL	583
DB	565	EDQLTKWGDFAQFVDKCKCAADKDCNFATGEPNLVARSKEAL	607
RESULT 10			
ALBU	PITG	STANDARD;	PRT; 605 AA.
ID	ALBU_PITG	STANDARD;	PRT; 605 AA.
AC	P08835; Q29018;		
DT	01-NOV-1988 (Rel. 09, Created)		
DT	01-NOV-1988 (Rel. 09, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Serum albumin precursor (Fragment).		
GN	ALB.		
OS	Sus scrofa (Pig).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
ON	NCBI_TaxID=9823;		
RX	[1]		
RP	SEQUENCE FROM N.A.		
RX	TISSUE=Liver;		
RX	MEDLINE=89016582; PubMed=3174440;		
RA	Baldwin G.S., Weinstein J.;		
RT	*Nucleotide sequence of porcine liver albumin.*;		

RL Nucleic Acids Res. 16:9045-9045(1988).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X12422; CAA30970.1; -  
 CC EMBL; M36787; AAA30988.1; -  
 CC PIR; S01382; ABPGS.  
 CC HSP; P02768; 1B7H.  
 CC InterPro: IPR000264; Serum\_albumin.  
 CC Pfam: PF00273; transport\_prot; 3.  
 CC PRINTS; PR00802; SERUMALBUMIN.  
 CC PRODOM; PD002486; Serum\_albumin; 1.  
 CC SMART; SM00103; ALBUMIN; 3.  
 CC PROSITE; PS00212; ALBUMIN; 3.  
 CC Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
 FT NON\_TER 1 1  
 FT SIGNAL <1 16 BY SIMILARITY.  
 FT PROPEP 17 22 BY SIMILARITY.  
 FT CHAIN 23 605 SERUM ALBUMIN.  
 FT DOMAIN 23 202 ALBUMIN 1.  
 FT DOMAIN 209 394 ALBUMIN 2.  
 FT DOMAIN 401 592 ALBUMIN 3.  
 FT METAL 31 31 COPPER (BY SIMILARITY).  
 FT DISULFID 75 84 BY SIMILARITY.  
 FT DISULFID 97 113 BY SIMILARITY.  
 FT DISULFID 112 123 BY SIMILARITY.  
 FT DISULFID 145 190 BY SIMILARITY.  
 FT DISULFID 189 198 BY SIMILARITY.  
 FT DISULFID 221 267 BY SIMILARITY.  
 FT DISULFID 266 274 BY SIMILARITY.  
 FT DISULFID 286 300 BY SIMILARITY.  
 FT DISULFID 299 310 BY SIMILARITY.  
 FT DISULFID 337 382 BY SIMILARITY.  
 FT DISULFID 381 390 BY SIMILARITY.  
 FT DISULFID 413 459 BY SIMILARITY.  
 FT DISULFID 458 469 BY SIMILARITY.  
 FT DISULFID 482 498 BY SIMILARITY.  
 FT DISULFID 497 508 BY SIMILARITY.  
 FT DISULFID 535 580 BY SIMILARITY.  
 FT DISULFID 579 588 BY SIMILARITY.  
 FT CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).  
 SQ SEQUENCE 605 AA; 69410 MW; 3E556BDD0D1A1FF CRC64;  
 Query Match 77.7%; Score 2411.5; DB 1; Length 605;  
 Best Local Similarity 76.0%; Pred. No. 1.4e-149;  
 Matches 438; Conservative 67; Mismatches 70; Indels 1; Gaps 1;  
 QY 1 DAHSEVAHFKDLDGENFALVLIAPQYLOQCPFDHVKLVNVTETKTCVADSEAE 60  
 DB 23 DTYKSEIAHFKDLDGEQYFGLVLIASFQHQCPYEHVKLVNVTETKTCVADSEAE 82  
 QY 61 NCDKSHUFLGDKLCTVATREYNGEMADCAKOPERNCECTLOHKKDDPNLPRLVPEV 120  
 DB 83 NCDKSHUFLGDKLCAIPSRREYHGLADCCKEPERNECFLOHNDPNLPKL-KDPP 141  
 QY 121 DVMCYAFHNDNEETFKLYLYELIARRHPYFYAPPELLFTFAKRYKAAFTCCQADKAACLLP 180  
 DB 142 VALCADQFQDEQKFWGKLYELIARRHPYFYAPPELLYVATIKYDVFSECCQAADKAACLLP 201

QY 181 KLDELDEGKASSAKORLKCASIQKQGERAFKAWAVARLSQRPKAEFAEYVKLVTDLT 240  
 DB 202 KIEHLREKVTSAKORLKCASIQKQGERAFKAWAVARLSQRPKAEFAEYVKLVTDLT 261  
 QY 241 VTECHGDLLECADRADLAKYICENODSISSSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
 DB 262 VHKECHGDLLECADRADLAKYICENODTISTIKLEKCEKPLLEKSHCIAEAKRDELPA 321  
 QY 301 DLPSLAADFVESKDYCKNAEAKDYFLGMFLYFYARHPDYSVLLRLAKTYVTFLEK 360  
 DB 322 DNLPEHDFVDEKVKCKNKEAKDYFLGTFYFESRHDPDYSVLLRLAKTYVTFLEK 381  
 QY 361 CAAADPHECYAKVDFEPLVEEPQNLKQNCLEFQOLGEYFQNALVRYTKVQPVST 420  
 DB 382 CAKEDPPACYATVDFKQPLVDEPKNLKQNCLEFQOLGEYFQNALVRYTKVQPVST 441  
 QY 421 PTLVEVSRNLKGVSKCKCKPFAKMPCAEDYLSVLNQLCVLHETKTPVSDRVTKCCTES 480  
 DB 442 PTLVEVARKLGLVSGRCCKPEERLSCAEDYLSVLNQLCVLHETKTPVSKVTKCCTES 501  
 QY 481 LVNRRPCFSALEVDVETVVPKEFNAETFTFHADICTLSEKQIKKOTALVELVHKPKAT 540  
 DB 502 LVNRRPCFSALTPTDETVPKPEFVEGTFTFHADLCTLPEDKEQIKKOTALVELLKHKPHAT 561  
 QY 541 KEOLKAVMDFAAEVKECKKADDKETCFAEKGGKLV 576  
 DB 562 EQLRTVLGNFAAFVQKCAAPDHEACFAVGPFRV 597  
 RESULT 11  
 ALBU\_MERUN  
 ID ALBU\_MERUN STANDARD; PRT; 609 AA.  
 AC Q35090;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Serum albumin precursor.  
 GN ALB.  
 OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).  
 OC Eukaryota; Metazoa; Chordata; Gratiata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;  
 OC Meriones.  
 OC NCBI\_TaxID=10047;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MGS IDR; TISSUE=Liver;  
 RX MEDLINE=98116663; PubMed=9455485;  
 RA Yoshida K., Seto-Ohshima A., Sinohara H.;  
 RT "Sequencing of cDNA encoding serum albumin and its extrahepatic  
 RL synthesis in the Mongolian gerbil, Meriones unguiculatus.";  
 RL DNA Res. 4:351-354(1997).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AB006197; BAA21765.1; -  
 CC PIR; JC5838; JC5838.  
 CC HSP; P02768; 1B7B.  
 CC InterPro: IPR000264; Serum\_albumin.  
 CC Pfam: PF00273; transport\_prot; 3.





CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X60688; CAA43098.1; -;  
 DR EMBL; V00381; CAA23680.1; -;  
 DR PIR; S15571; ABCHS.  
 DR HSP; P02768; IE7B.  
 DR InterPro; IPR000264; Serum\_albumin.  
 DR Pfam; PF00273; transport\_prot; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum\_albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 DR Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
 FT SIGNAL 1 18  
 FT PROPEP 19 23  
 FT CHAIN 24 615 SERUM ALBUMIN.  
 FT DOMAIN 24 209 ALBUMIN 1.  
 FT DOMAIN 216 401 ALBUMIN 2.  
 FT DOMAIN 408 599 ALBUMIN 3.  
 FT METAL 30 30 COPPER (BY SIMILARITY).  
 FT DISULFID 80 89 BY SIMILARITY.  
 FT DISULFID 102 118 BY SIMILARITY.  
 FT DISULFID 117 128 BY SIMILARITY.  
 FT DISULFID 152 197 BY SIMILARITY.  
 FT DISULFID 196 205 BY SIMILARITY.  
 FT DISULFID 228 274 BY SIMILARITY.  
 FT DISULFID 273 281 BY SIMILARITY.  
 FT DISULFID 293 307 BY SIMILARITY.  
 FT DISULFID 306 317 BY SIMILARITY.  
 FT DISULFID 344 389 BY SIMILARITY.  
 FT DISULFID 388 397 BY SIMILARITY.  
 FT DISULFID 420 466 BY SIMILARITY.  
 FT DISULFID 465 476 BY SIMILARITY.  
 FT DISULFID 489 505 BY SIMILARITY.  
 FT DISULFID 504 515 BY SIMILARITY.  
 FT DISULFID 542 587 BY SIMILARITY.  
 FT DISULFID 586 595 BY SIMILARITY.  
 FT CARBOHYD 500 500 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CONFLICT 24 24 F -> M (IN REF. 3).  
 SQ SEQUENCE 615 AA; 69918 MW; E59E4BCEC066C6 CRC64;  
 Query Match 50.2%; Score 1557.5; DB 1; Length 615;  
 Best Local Similarity 46.7%; Pred. No. 4.3e-94;  
 Matches 273; Conservative 118; Mismatches 192; Indels 1; Gaps 1;  
 QY 3 HKSEVAHPRKDLGNEKALVLIATFAQLQCPEDHVKLVNEVTEFAKTCVADSEANBC 62  
 DB 30 HKSEIAHRYNDLKEETEFKAVAMITFAQLQRCSEGLSKLVQDVLDLQCKVANEDEPEC 89  
 QY 63 DKSLLHTLGDKLCVATLRETYGEMADCAKQEPNECBFLQHKDDNPNLPR-LVPRVD 121  
 DB 90 SKFLPSIIILDCQVEKLRDSYGAMADCCSKADPERNECTLSFKVQDPFQYQRPASD 149  
 QY 122 VMCTAFHNEFTFKLYLFIARHPHYFAPLFFPAKRYKAATFCCQAADKAACLXPX 181  
 DB 150 VICQEQDNRVSVLGHETYSVARRHPFLYAPALLSFAYDFEHALQSCCKESDVGACLDTK 209  
 QY 182 LDELDEKASAKORLCKASLOKFGERAKAWARLSQRFPAFAEVSRLVLDLTKY 241  
 DB 210 EIVMREKAGVSVQQTFCGLLKQFGDRVQARQIYLSQKPKAPFSEVSFVHDSIGV 269  
 QY 242 HFECHGOLLECADRADLAKYICENQDSISKKECEKPKLLEKSHCHIAEVENDEPAD 301  
 DB 270 HRECCGDVCEMDMAMNSNLCSSQDVFSGKINDCCERPIVERSQCIWEAFDEKPAD 329  
 QY 302 LPSLAADFVESKDVCKNFAKADKDFLGMFLFYEARHPDYVSWLLRLAKTYETLEKCC 361  
 DB 330 LPSLVEKIEDKEVCKSFAGHDAPFAEIVFYSRHPFTSIQLIMRIAKGYESLLEKCC 389

QY 362 AADPHECYAKVDFEKFPLVEPQNLIKQNCLEFQEGYKFNQALLVRYTKKVPQVSTP 421  
 DB 390 KTDNPAECYANAOEQUNQHIEKTOBVVYKNTCDLHDHGEADFLKSILIRTKKMPQVPTD 449  
 QY 422 TLVSVSNLKGVSCKCKHPAKMPCAEYDLSVLNQLCVLHEKTPVSDRVTKCCTESTL 481  
 DB 450 LLLTGGKMTTIGTKCCQIGEDRMACEGYLSIVHDTCRKOETTPINDVNSQCCSLY 509  
 QY 482 VNRPCPSALEVDYVYKFNNAETFTTHADICTLSEKERQIKKQKTALVELVHKHAKK 541  
 DB 510 ANRPPCTAMGVDTKYVPPFPNDFMSFDEKLSAFABEREVGQKLLINLIRKPKQMT 569  
 QY 542 EQLKAVMDFAFVEKCKKADDKETCFABEKKKLVAAQALGL 585  
 DB 570 EQIKTIADGFTAMVCKCKQSDINFCFGEAGANLIVOSRHILGI 613  
 RESULT 14  
 ID PETA\_PANTR STANDARD; PRG; 609 AA.  
 AC Q28789;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-fetoprotein).  
 DE fetoprotein).  
 GN AFP.  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
 OX NCBI\_TaxID=9598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96032345; PubMed=7557431;  
 RA Nishio H., Gibbs P.E., Minghetti P.P., Zielinski R., Dugaiczak A.;  
 RT "The chimpanzee alpha-fetoprotein-encoding gene shows structural  
 RT similarity to that of gorilla but distinct differences from that of  
 RT human";  
 RL Gene 162:213-220(1995).  
 CC -I- FUNCTION: BINDS COPPER, NICKEL, AND FATTY ACIDS AS WELL AS, AND  
 CC BILIRUBIN LESS WELL THAN, SERUM ALBUMIN. ONLY A SMALL PERCENTAGE  
 CC (LESS THAN 2%) OF THE HUMAN AFP SHOWS ESTROGEN-BINDING PROPERTIES.  
 CC -I- SUBUNIT: DIMERIC AND TRIMERIC FORMS HAVE BEEN FOUND IN ADDITION  
 CC TO THE MONOMERIC FORM.  
 CC -I- SUBCELLULAR LOCATION: Secreted.  
 CC -I- TISSUE SPECIFICITY: Plasma. Synthesized by the fetal liver and  
 CC yolk sac.  
 CC -I- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.  
 CC -I- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U21916; AAA91641.1; -;  
 DR PIR; JC4258; JC4258.  
 DR HSP; P02768; IE7B.  
 DR InterPro; IPR000264; Serum\_albumin.  
 DR Pfam; PF00273; transport\_prot; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum\_albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 2.  
 DR Glycoprotein; Sulfation; Repeat; Metal-binding; Copper; Nickel;  
 KW Signal.  
 FT SIGNAL 1 18 BY SIMILARITY.  
 FT CHAIN 19 609 ALPHA-FETOPROTEIN.  
 FT DOMAIN 20 205 ALBUMIN 1.  
 FT DOMAIN 212 397 ALBUMIN 2.

```

FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 22 22 COPPER AND NICKEL (BY SIMILARITY).
FT DISULFID 99 114 BY SIMILARITY.
FT DISULFID 113 124 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 609 AA; 68741 MW; C032987CADOE672B CRC64;

Query Match 40.4%; Score 1253.5; DB 1; Length 609;
Best Local Similarity 40.1%; Pred. No. 2.4e-74;
Matches 236; Conservative 115; Mismatches 231; Indels 7; Gaps 3;

QY 3 HKSE-----VAHREKDLGEENEKALVLIAPAQVLOQCPEHVDKLYNVEFFAKTCVADR 57
DB 22 HRNEYGIASLDSTQCTACINLTDLATIFFAQVQVATYKVKMKVDALTAIEKPTGDE 81
QY 58 SAENCKSLHTFGKLCVAVLRTETGEMADCCAQEPERNECFLOKHQDNP-NLPRLV 116
DB 82 QSAGLENQLPATLEBCEKEKLEKYGH-SDCCSQSEGRHNCFLAHKPTPASIFPQ 140
QY 117 RPEVDVMTAFHNETFLKYLVEIARRHPYFAPPELLFFAKRYKAFAFTCCOAKDKAA 176
DB 141 VPPEVTCEAYEDRETFNFKTYEIAARRHPFTYAPTILLWAARYDKLIPSCCKAENAYE 200
QY 177 CLLPKLDLDEKASKAKORLKCASQKFCGERAFKAWAVARLSORFFKAFBVSRLVT 236
DB 201 CEQYKATVYKELRESLLNHACAVMKNGTFTFOATVTKLSOKFKVNFTEIQKLVL 260
QY 237 DLTKVHTCCGDLBCADRADLAKYICENQDISSKLKECCCKEPLKSHCHIAEVND 296
DB 261 DVAVHHECCRGVDLQDGEKIMSYTCQQDTLSNKTCECKLTITLQGCIIHAEND 320
QY 297 EMPADLPSLAADVESKVCNKYAEKDVPLGMELEYARRHPDYSVVLLRLAKTYET 356
DB 321 EKPGLSPNLRCTGDRDNQFSSEKRNIFLASVFEISRRPQLANSVILRVANGQEL 380
QY 357 LEKCCAAADPHCYAKVDFEKPVLVEEPQNLIKQNCSELFQELGEYKFFQNALLYRTKKVP 416
DB 381 LEKCFQTEPLEQDQGEELQYIQESQALAKRSCLGPKLGEYLLQNALVAYTKKAP 440
QY 417 QYSTPTLVEVSNLKGSKCKHPEAKMEPCADLYSVVNLQVLHEKTPSDRYTKC 476
DB 441 QJTSSELMAITRKWATAATACQJSEDKLLACGEGAADIIGHLCIRHETTPVNGVQC 500
QY 477 CTESIVNRPCFSALVDETYVPKEFNAETFTFADICTLSKEKQIKKQATALVELYKHK 536
DB 501 CTSSYANRPFCSLLVDETYVPFASDDKFIHFKDLCOAQGVALTMTKQELINLYKQK 560
QY 537 PKATKEQLKAVMDFAAFVEKCKKADDKETCFABEGKKLVASQALGI 585
DB 561 PQITBEQLEAVIADFSLLEKCCQGEQVCFABEGOKLISKTRALGV 609

```

RESULT 15

PETA\_HUMAN

ID PETA\_HUMAN STANDARD; PRT; 609 AA.

AC P02771;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-
DE fetoprotein).
DE GN AFF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=83273664; PubMed=6192439;
RX Morinaga T., Sakai M., Wegmann T.G., Tamaoki T.;
RT "Primary structures of human alpha-fetoprotein and its mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:4604-4608(1983).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=87185438; PubMed=2436661;
RX Gibbs P.E.M., Zielinski R., Boyd C., Dugaiczky A.;
RT "Structure, polymorphism, and novel repeated DNA elements revealed by
a complete sequence of the human alpha-fetoprotein gene.";
RL Biochemistry 26:1332-1343(1987).
RN [3]
RN SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=93278385; PubMed=7684942;
RX McVey J.H., Michaelides K., Hansen L.P., Ferguson-Smith M.,
RA Tilghman S., Krumlauf R., Tuddenham E.G.D.;
RT "A G->A substitution in an HNF I binding site in the human alpha-
fetoprotein gene is associated with hereditary persistence of alpha-
fetoprotein (HPAFP).";
RL Hum. Mol. Genet. 2:379-379(1993).
RN [4]
RN SEQUENCE OF 429-556 FROM N.A.
RX MEDLINE=83158778; PubMed=6187626;
RX Beattie W.G., Dugaiczky A.;
RT "Structure and evolution of human alpha-fetoprotein deduced from
partial sequence of cloned cDNA.";
RL Gene 20:415-422(1982).
RN [5]
RN PARTIAL SEQUENCE OF 19-609.
RX MEDLINE=91242409; PubMed=1709810;
RX Pucci P., Scilliano R., Malorni A., Marino G., Tecce M.F.,
RA Caccarini C., Ferrana B.;
RT "Human alpha-fetoprotein primary structure: a mass spectrometric
study.";
RL Biochemistry 30:5061-5066(1991).
RN [6]
RN PRELIMINARY SEQUENCE OF 19-35.
RX MEDLINE=77242506; PubMed=70228;
RX Tachini S., Hsu K., Heinrichson R.L., Miller J.B.;
RT "Studies on human alpha-fetoprotein. Isolation and characterization
of monomeric and polymeric forms and amino-terminal sequence
analysis.";
RL Biochim. Biophys. Acta 493:418-428(1977).
RN [7]
RN PRELIMINARY SEQUENCE OF 19-38.
RX MEDLINE=78001760; PubMed=71198;
RX Aoyagi Y., Ikenaka T., Ichida F.;
RT "Comparative chemical structures of human alpha-fetoproteins from
fetal serum and from ascites fluid of a patient with hepatoma.";
Cancer Res. 37:3663-3667(1977).
RN [8]
RN PRELIMINARY SEQUENCE OF 19-39.
RX MEDLINE=75018719; PubMed=4138095;
RX Ruoslahti E., Pihko H., Vaheri A., Seppala M., Virolainen M.,
RA Kontinen A.;
RT "Alpha fetoprotein: structure and expression in man and inbred mouse
strains under normal conditions and liver injury.";
RL Johns Hopkins Med. J. Suppl. 3:249-255(1974).
RN [9]
RN GENE STRUCTURE.
RX MEDLINE=85182629; PubMed=2580830;
RX Sakai M., Morinaga T., Urano Y., Watanabe K., Wegmann T.G.,
RA Tamaoki T.;
RT "The human alpha-fetoprotein gene. Sequence organization and the 5'

```

flanking region.";  
 J. Biol. Chem. 260:5055-5060(1985).  
 [10]  
 RN METAL-BINDING.  
 RP MEDLINE=79001617; PubMed=80265;  
 RA Aoyagi Y., Ikenaka T., Ichida F.;  
 "Copper (II)-binding ability of human alpha-fetoprotein.";  
 Cancer Res. 38:3483-3486(1978).  
 [11]  
 RN BILIRUBIN-BINDING.  
 RP MEDLINE=80001710; PubMed=89900;  
 RA Aoyagi Y., Ikenaka T., Ichida F.;  
 "Alpha-Fetoprotein as a carrier protein in plasma and its bilirubin-binding ability.";  
 Cancer Res. 39:3571-3574(1979).  
 [12]  
 RN SULFATION.  
 RP MEDLINE=86042625; PubMed=2414772;  
 RA Liu M.C., Yu S., Sy J., Redman C.M., Lipmann F.;  
 "Tyrosine sulfation of proteins from the human hepatoma cell line HepG2.";  
 Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).  
 CC -!- FUNCTION: BINDS COPPER, NICKEL, AND FATTY ACIDS AS WELL AS, AND BILIRUBIN LESS WELL THAN, SERUM ALBUMIN. ONLY A SMALL PERCENTAGE (LESS THAN 2%) OF THE HUMAN AFP SHOWS ESTROGEN-BINDING PROPERTIES.  
 CC -!- SUBUNIT: DIMERIC AND TRIMERIC FORMS HAVE BEEN FOUND IN ADDITION TO THE MONOMERIC FORM.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: PLASMA. SYNTHESIZED BY THE FETAL LIVER AND YOLK SAC.  
 CC -!- DEVELOPMENTAL STAGE: OCCURS IN THE PLASMA OF FETUSES MORE THAN 4 WEEKS OLD, REACHES THE HIGHEST LEVELS DURING THE 12TH-16TH WEEK OF GESTATION, AND DROPS TO TRACE AMOUNTS AFTER BIRTH. THE SERUM LEVEL IN ADULTS IS USUALLY LESS THAN 40 NG/ML. AFP OCCURS ALSO AT HIGH LEVELS IN THE PLASMA AND ASCITIC FLUID OF ADULTS WITH HEPATOMA.  
 CC -!- PTM: INDEPENDENT STUDIES SUGGEST HETEROGENEITY OF THE AMINO-TERMINAL SEQUENCE OF THE MATURE PROTEIN AND OF THE CLEAVAGE SITE OF THE SIGNAL SEQUENCE.  
 CC -!- PTM: SULFATED.  
 CC -!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 -----  
 DR EMBL; M10949; AAA51674.1; -;  
 DR EMBL; M10950; AAA51675.1; -;  
 DR EMBL; V01514; CAA24758.1; -;  
 DR EMBL; M16110; AAB58754.1; -;  
 DR EMBL; Z19532; CAA79592.1; -;  
 DR PIR; A26624; FPHU.  
 DR KSSP; P02768; IE7B.  
 DR GlycoSuiteDB; P02771; -;  
 DR Siena-2DPAGE; P02771; -;  
 DR Genew; HGNC:317; AFP.  
 DR MTM; 104150; -;  
 DR InterPro; IPR000264; Serum\_albumin.  
 DR Pfam; PF00273; transport\_prot; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum\_albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 2.  
 KW Glycoprotein; Sulfation; Repeat; Metal-binding; Copper; Nickel;  
 KW Signal; Polymorphism.  
 FT SIGNAL 1 18  
 FT CHAIN 19 609 ALPHA-FETOPROTEIN.  
 FT DOMAIN 20 205 ALBUMIN 1.  
 FT DOMAIN 212 397 ALBUMIN 2.

FT	DOMAIN	404	595	ALBUMIN 3.
FT METAL	22	99	114	COPPER AND NICKEL.
FT DISULFID	113	124	193	
FT DISULFID	192	201	277	
FT DISULFID	269	277	303	
FT DISULFID	289	302	313	
FT DISULFID	384	393	462	
FT DISULFID	416	472	501	
FT DISULFID	485	511	583	
FT DISULFID	538	582	591	
FT CARBOHYD	251	251	251	
FT VARIANT	570	570	570	
FT FTIG-VAR_012049.				
SQ SEQUENCE	609 AA; 68677 MW; 4D4E45820E1C2D4F CRC64;			
Query Match	40.3%; Score 1249.5; DB 1; Length 609;			
Best Local Similarity	39.3%; Pred. No. 4.4e-74;			
Matches	235; Conservative 116; Mismatches 231; Indels 7; Gaps 3;			
QY	3 HKSE-----VAHFRKDLGEENFKALVLAFAQYLQCCPPFEDHVKLVNEVTEFAKTCVADE 57			
DB	22 HNEVGTASILDYSVOCTAELSLADLATTFFAQQVQYATYKEVSKVMDALTALEKPTGDE 81			
QY	58 SAENCDLSLHTLFGDKLCVVAIRITYGEMADCAKOEKPERNECFLOHKDNP-NLPRLV 116			
DB	82 OSSGLENQLPAFTELECHEKEILEKYGH-SDCGSQSEGRHNCFLAHKPTPASILPQ 140			
QY	117 REVDVCTAFHNEETFLKLYLEIARRHEFFVAPPELLFFAKRYKAAATECCQAADKAA 176			
DB	141 VPEVTSCEAYEEDRETFMKNFIETIARRHFFLAPFILLWAARYDKIIPSCCKAENAVE 200			
QY	177 CLLPKLDELDEGKASSAKORLKCAKLOKGERAFKAWAVARLSORFPAEYVSKLVT 236			
DB	201 CFQTKAATVTKELRESSLLNQHACAVMKNFGRFQAITVTKLSQKFTVNETEQKLVL 260			
QY	237 DLTQVHTECHGDLLECADDRLADLAKYICENQDSISSKLECECEKPLEKSHCIAEVEND 296			
DB	261 DYAHVHECCRGDVLDCLOQDGKIMSYICSOQDTLSNKTETECCKLTTLERGQCIHAEND 320			
QY	297 EMPADLPSLAADFVESKDVCKNVAEAKDVFELGMELVEYARRHPDYSVLLLELAKTYETT 356			
DB	321 EKPEGLSPNRLRGDRDNQFSSGKKNIFLASVHEISRRHPQLAVSVILKRVANGYOEL 380			
QY	357 LEKCAAADPHECYAKVDFDEKPLVEEPQNLKONCELPQLGEYKFQNALIVRYTKVP 416			
DB	381 LEKCFQTEPLECKDGKEELQYIQESQALAKSKSGLFQKLGYYLQNAFLVATYTKAP 440			
QY	417 QVSTPLTVESRNLGKVGSKCKHPEAKMPCADYLSVVLNOLCVLHEKTPVSDRVTKC 476			
DB	441 QLTSSSELMAITRKMAATAATACQQLSEDKILLACGEGAADIIGHICLRHHEMTVPNPGVQC 500			
QY	477 CTESLVNRRPCFSALVEDETYVPKEFNATETTFHADICTLSEKEROIKKOTALVELYKHK 536			
DB	501 CTSSVANRRPCFSSLVVDVETVYPPAFSDKFIHKDLCQAQGVALQTMKQFLINLVKQK 560			
QY	537 PRATKEQLKAVMDQFAAFVEKCKCKADDEKCTFAEEGKLVAAASQAALGL 585			
DB	561 POLTEQLKAVIADFSGLLEKCCQCEQVEVCFAGEGKLSIKYTRALGV 609			

Search completed: August 31, 2003, 16:33:47  
 Job time : 28 secs



GenCore version 5.1.1.6

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2003, 16:33:10 ; Search time 106 Seconds

(without alignments)  
1424.158 Million cell updates/sec

Title: US-09-833-041-18

Perfect score: 3103

Sequence: 1 DAHKEVAHRFKDLGRENFK.....TCFAEKGKLVAAQAALGL 585

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23.\*

1: sp.archaea.\*

2: sp.bacteria.\*

3: sp.fungi.\*

4: sp.human.\*

5: sp.invertebrate.\*

6: sp.mammal.\*

7: sp.mhc.\*

8: sp.organella.\*

9: sp.phase.\*

10: sp.plant.\*

11: sp.rodent.\*

12: sp.virus.\*

13: sp.vertebrate.\*

14: sp.unclassified.\*

15: sp.rvirus.\*

16: sp.bacteriap.\*

17: sp.bacteriap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2504	80.7	608	5 Q95VB7	Q95vb7 schistosoma
2	2374	76.5	608	11 Q8C7H3	Q8c7h3 mus musculus
3	2336	75.3	576	11 Q8C7C7	Q8c7c7 mus musculus
4	1865.5	60.1	396	4 Q8IUK7	Q8iuk7 homo sapien
5	1295.5	41.7	527	13 Q8JIA9	Q8jia9 sphendonod p
6	1242	40.0	609	6 Q8MJU5	Q8mju5 canis famil
7	1242	40.0	626	13 Q8UW05	Q8uw05 ambystoma m
8	1218.5	39.3	610	6 Q8MJ76	Q8mj76 sus scrofa
9	1087	35.0	624	13 Q8UW06	Q8uw06 ambystoma t
10	1082	34.9	605	11 Q8BK65	Q8bk65 mus musculus
11	1076	34.7	605	11 Q8BK56	Q8bk56 mus musculus
12	1045	33.7	400	13 Q8JIA7	Q8jia7 sphendonod p
13	955	30.8	603	13 Q8YGH6	Q8ygh6 rana shqipe
14	928.5	29.9	614	13 Q91134	Q91134 najja najja
15	888	28.6	406	13 Q8JIA8	Q8jia8 hoplodactyl
16	739	23.8	205	11 Q8CG74	Q8cg74 mus musculus

17	713	23.0	417	11 Q8R0J9	Q8r0j9 mus musculu
18	376.5	12.1	484	13 Q9W6F5	Q9w6f5 gallus gall
19	373	12.0	476	11 Q9CY31	Q9cy31 mus musculu
20	372	12.0	476	11 Q9IXG1	Q9ixg1 mus musculu
21	365	11.8	551	13 Q42279	Q42279 petromyzon
22	331	10.7	122	13 Q9WZ8	Q9wz8 larus argen
23	290	9.3	123	13 Q9WZ6	Q9wz6 poephila gu
24	264	8.5	135	11 Q63205	Q63205 rattus norv
25	188	6.1	1723	2 Q9JMK8	Q9jmx8 helicobacte
26	184	5.9	1819	16 Q9LIV0	Q9livo helicobacte
27	184	5.9	1927	16 Q25262	Q25262 helicobacte
28	162.5	5.2	44	6 Q95MC2	Q95mc2 equus cabal
29	162.5	5.2	680	5 Q9V6S8	Q9v6s8 drosophila
30	161	5.2	1079	3 Q96V11	Q96v11 pneumocysti
31	156	5.0	1026	3 Q74669	Q74669 pneumocysti
32	154.5	5.0	3843	5 Q9USD0	Q9u5d0 drosophila
33	153.5	4.9	661	5 Q8MS79	Q8ms79 drosophila
34	153.5	4.9	3843	5 Q9VT94	Q9vu94 drosophila
35	153	4.9	62	6 Q8M111	Q8m111 sus scrofa
36	150	4.8	1065	3 Q01828	Q01828 pneumocysti
37	149	4.8	40	6 Q9TRA5	Q9tra5 oryctolagus
38	148	4.8	1028	3 Q74668	Q74668 pneumocysti
39	144.5	4.7	8749	4 Q8NF91	Q8nf91 homo sapien
40	142.5	4.6	2756	10 Q9LJ60	Q9lj60 arabidopsis
41	141.5	4.6	1560	5 Q26644	Q26644 strongyloce
42	138.5	4.5	1069	3 Q96VI2	Q96vi2 pneumocysti
43	137.5	4.4	1348	16 Q8YK55	Q8yk55 anabaena sp
44	135.5	4.4	2841	5 Q8MLU9	Q8mlu9 drosophila
45	135.5	4.4	2931	5 Q9W2C6	Q9w2c6 drosophila

## ALIGNMENTS

RESULT 1

Q95VB7 PRELIMINARY; PRT; 608 AA.

AC Q95VB7;

DT 01-DEC-2001 (TREMREL. 19, Created)

DT 01-DEC-2001 (TREMREL. 19, Last sequence update)

DT 01-MAR-2003 (TREMREL. 23, Last annotation update)

DE Albumin.

OS Schistosoma mansoni (Blood fluke).

OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeiida;

OC Schistosomatidae; Schistosomatidae; Schistosoma.

CX NCBI\_taxid=6183;

RN [1]

RP SEQUENCE FROM N.A.

RA Osman A., Asahi H., Staecker M.J., LoVerde P.T.;

RT "Albumin precursor homolog is a novel T helper cell immunogenic egg

RT component in murine infection with Schistosoma mansoni.,"

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF418550; AAL08579.1; ..

DR InterPro; IPR000264; Serum\_albumin.

DR Pfam; PF00273; transport\_prot; 3.

DR PRINTS; PRO0802; SERUMALBUMIN.

DR ProDom; PD002486; Serum\_albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 2.

DR SEQUENCE 608 AA; 68225 MW; E5EABB28E1C66E54 CRC64;

Query Match 80.7%; Score 2504; DB 5; Length 608;

Best Local Similarity 76.3%; Pred. No. 1.7e-188;

Matches 445; Conservative 79; Mismatches 59; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDLGRENFKALVIAFAQYLOQCFEDHVLVNEVTEFAKTCVADESAAE 60

Db 25 DAHKEVAHRFKDLGRENFKALVIAFAQYLOQCFEDHVLVNEVTEFAKTCVADESAAE 84

QY 61 NCKSLHTLFGKLCIAITFLRTYTGEMADCCAKQPERNECFLOHKDDNINLPLVRPV 120

Db 85 NCKSLHTLFGKLCIAITFLRTYTGEMADCCAKQPERNECFLOHKDDNINLPLVRPV 144





```

QY 249 DLLECAADRADLAKYICENQDSISSKLECCCKPPLLEKSHCIAVENDEMPADLPISLAAD 308
Db 241 DLLECAADRAELAKYICENQATISSKLOTCCDKPLKKAHLSEVHDTPADLPISLAAD 300
QY 309 FVSKDVCKNYABADKDFLGMFLYEVARRHPDYSVLLLRLLAKTYETTLKCCCAADPHE 368
Db 301 FVEQEVCKNYABADKDFLGMFLYEVARRHPDYSVLLLRLLAKTYETTLKCCCAADPHE 360
QY 369 CYAKVDFERKPLVEEPQNLKONCELFQELGEYKFNALLVRYTKKYPQVSTPTLVEVSR 428
Db 361 CYGVVLAEPQPLVEEPKVLNVCNLYEKLGEYGFQNALVRYTKKYPQVSTPTLVEAR 420
QY 429 NLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHTPVSVDRTVTKCCTESLVNRPQF 488
Db 421 NLGRVTKCTLPEDQRLPCVEDYLSAILNRVCLLHEKTPVSEHVTCCSGSLVRRPQF 480
QY 489 SALEVDETVVPKEFNAETFTPHADICVLSKERQIKKQTAVALVELVKKPKATKEQLKATM 548
Db 481 SALTVDETVVPKEFNAETFTPHADICVLSKERQIKKQTAVALVELVKKPKATKEQLKATM 540
QY 549 DDPAAVEKCCADKDKETCFABEGKKLVAASQAAL 583
Db 541 DDPAAVEKCCADKDKETCFABEGKKLVAASQAAL 575

RESULT 4
Q8UJ7 PRELIMINARY; PRT; 396 AA.
ID Q8UJ7
AC Q8UJ7
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE Similar to serum albumin precursor.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC035969; AA035969.1;
SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

```

```

Query Match 60.1%; Score 1865.5; DB 4; Length 396;
Best Local Similarity 63.6%; Pred. No. 1.6e-138;
Matches 372; Conservative 0; Mismatches 0; Indels 213; Gaps 1;

QY 1 DAHSEVAFHRFDLGEENFKALVLIAPAFQYLOQCPEDHVKLVNVEVTEFAKTCVADSEAE 60
Db 25 DAHSEVAFHRFDLGEENFKALVLIAPAFQYLOQCPEDHVKLVNVEVTEFAKTCVADSEAE 84
QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVPRV 120
Db 85 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVPRV 144
QY 121 DVMTAFPHNEETFLKKYLIEARRHPYFYAPPELLFAKRYKAATECCQAADKACLP 180
Db 145 DVMTAFPHNEETFLKKYLIEARRHPYFYAPPELLFAKRYKAATECCQAADKACLP 163
QY 181 KLDELDEGKASSAKQRLKASLQKGFGERAFKAWAVARLSQRPKAEFVSKLVDTLK 240
Db 164 ----- 163
QY 241 VHTCCGDLLECAADRADLAKYICENQDSISSKLECCCKPPLLEKSHCIAVENDEMPA 300
Db 164 ----- 163
QY 301 DLPSLAADFVSKDVCKNYABADKDFLGMFLYEVARRHPDYSVLLLRLLAKTYETTLK 360
Db 164 ----- 163
QY 361 DLPSLAADFVSKDVCKNYABADKDFLGMFLYEVARRHPDYSVLLLRLLAKTYETTLK 360
Db 164 ----- 163

```

```

QY 361 CAADAPHECYAKVDFERKPLVEEPQNLKONCELFQELGEYKFNALLVRYTKKYPQVST 420
Db 172 CAADAPHECYAKVDFERKPLVEEPQNLKONCELFQELGEYKFNALLVRYTKKYPQVST 231
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHTPVSVDRTVTKCCTES 480
Db 232 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHTPVSVDRTVTKCCTES 291
QY 481 LVNRRPCFSALEVDETVVPKEFNAETFTPHADICVLSKERQIKKQTAVALVELVKKPKAT 540
Db 292 LVNRRPCFSALEVDETVVPKEFNAETFTPHADICVLSKERQIKKQTAVALVELVKKPKAT 351
QY 541 KEOLKAVMDPFAAFVEKCCADKDKETCFABEGKKLVAASQAALGL 585
Db 352 KEOLKAVMDPFAAFVEKCCADKDKETCFABEGKKLVAASQAALGL 396

RESULT 5
Q8UJ9 PRELIMINARY; PRT; 527 AA.
ID Q8UJ9
AC Q8UJ9
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Serum albumin (Fragment).
OS Sphenodon punctatus (Hatteria) (Tuatara).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.
OX NCBI_TaxID=8508;
RN [1]
RP SEQUENCE FROM N.A.
RA Metcalf V.J.; Brennan S.O.; George P.M.; Chambers G.K.;
RT "Partial coding sequence for Sphenodon punctatus 68 kDa albumin.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF375971; AA046104.1;
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD082486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
FT NON_TER 1
SQ SEQUENCE 527 AA; 59711 MW; C62B799E387F5929 CRC64;

```

```

Query Match 41.7%; Score 1295.5; DB 13; Length 527;
Best Local Similarity 45.9%; Pred. No. 1.4e-93;
Matches 240; Conservative 92; Mismatches 188; Indels 3; Gaps 3;

QY 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVPRV 120
Db 5 CLASLDTLFLDLCHIEGFAKY-DLAACCAKAEVERECLAHKNAATPGFIFAFORPGI 63
QY 121 DVMTAFPHNEETFLKKYLIEARRHPYFYAPPELLFAKRYKAATECCQAADKACLP 180
Db 64 EVSKLYQDRLTGLNLTIEVARRHPYLVQVPPVFATASLYDEALKTCCCTADKATCFP 123
QY 181 KLDELDEGKASSAKQRLKASLQKGFGERAFKAWAVARLSQRPKAEFVSKLVDTLK 240
Db 124 RIPLLEYLKMNGIQENWTCGILKFGERTLAKTLVQMSQKFPKADFAINKLVEDITH 183
QY 241 VHTCCGDLLECAADRADLAKYICENQDSISSKLECCCKPPLLEKSHCIAVENDEMPA 300
Db 184 MHTECCGDTLLECDRLALTEYTCSHKDAISSKLPCTCEKSVLEGECEVLENDREPA 243
QY 301 DLPSLAADFVSKDVCKNYABADKDFLGMFLYEVARRHPDYSVLLLRLLAKTYETTLK 360
Db 244 DLSEIAEVIEDPHVCDHLAKEDDAFLAKTUYEISRRHPDLSQILLGVCKGQELERC 303
QY 361 CAADAPHECYAKVDFERKPLVEEPQNLKONCELFQELGEYKFNALLVRYTKKYPQVST 420
Db 304 CKTDNPECYGAAEDLKKHIAQFQELVQOQNDLNTLGGYLFHALLIRYTRMPQLTS 363

```

QY	421	PTLYEVSRLNGKVGSKCKHPAKRMPCAEDYLNVINOLCVLHEKTPVSDRVTKCCTBS	480
Db	364	BELIFYRR-ITRAASRCCEVSDKKLPCTEGYDFVLGICQHQRHSINVNWCQCNS	422
QY	481	LVNRRPCFSALEVDENIYPFKFNAFTTFHADICTLSSEKEROIKQTALVELVKHKPKAT	540
Db	423	YALRSLCITSLGGDEKFVIPFISSADLTFFHFDCHAAQDKLOERQQQMIVNLVKHKPNIT	482
QY	541	KEQLKAYMDFAAFVFKCKCKADDKETCFEAEGKKLVAAASQAAL	583
Db	483	KEQLQTVFGGFTKMTEKCCAKAEHDACFGEFGKLVAESQTAL	525
RESULT 6			
Q8MJU5			
ID	Q8MJU5	PRELIMINARY;	PRT; 609 AA.
AC	Q8MJU5;		
DT	01-OCT-2002 (TrEMBLrel. 22, Created)		
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	Alpha-fetoprotein.		
GN	APP.		
OS	Canis familiaris (Dog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
OX	NCBI_TaxID=9615;		
XP	[1]		
RN	SEQUENCE FROM N.A.		
RA	Furuichi M., Neo S., Hisasue M., Tsuchiya R., Watanabe M.,		
RA	Hashizaki K., Hisamatsu S., Yamada T.;		
RT	"Canine alpha-fetoprotein cDNA";		
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB089789; EAC07513.1; -.		
DR	InterPro; IPR000264; Serum_albumin.		
DR	Pfam; PF00273; transport_prot; 3.		
DR	PRINTS; PR00802; SERUMALBUMIN.		
DR	ProDom; PD02486; Serum_albumin; 1.		
DR	SMART; SM00103; ALBUMIN; 3.		
DR	PROSITE; PS00212; ALBUMIN; 2.		
SQ	SEQUENCE 609 AA; 68782 MW; BE4B8250C5AF2AF0 CRC64;		

RESULT 7	Q8UW05	PRELIMINARY;	PRT;	626 AA.
ID	Q8UW05			
AC	Q8UW05;			
DT	01-VAR-2002 (Tremblrel. 20, Created)			
DT	01-VAR-2002 (Tremblrel. 20, Last sequence update)			
DT	01-Oct-2002 (Tremblrel. 22, Last annotation update)			
DE	Serum albumin precursor.			
OS	ALB.			
GN	Ambystoma maculatum (spotted salamander).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Caudata; Salamandroides; Ambystomatidae;			
OC	Ambystoma.			
OX	NCBI_TaxID=43114;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RA	Haverfield E.V., Uzzell T., Spolsky C.M., Bazariseren B.;			
RT	"Serum albumin of the mole salamanders Ambystoma maculatum and Ambystoma texanum.";			
RL	Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.			
DR	EMBL; AF217183; AAL56646.1; -			
DR	InterPro; IPR000264; Serum_albumin.			
DR	Pfam; Pf00273; transport_prot; 3.			
DR	PRINTS; PS00802; SERUMALBUMIN			
DR	ProDom; PD002486; Serum_albumin; 1.			
DR	SMART; SM00103; ALBUMIN; 3.			
DR	PROSITE; PS00212; ALBUMIN; 2.			
KW	Signal.			
FT	FT SIGNAL	1	24	POTENTIAL.
FT	CHAIN	25	626	SERUM ALBUMIN.
SQ	SEQUENCE	626 AA;	70677 MW;	9D66F5F174AC23F CRC64;
Query Match 40.0%; Score 1242; DB 13; Length 626;				
Best Local Similarity 40.1%; Pred. No. 2.8e-89;				
Matches 237; Conservative 110; Mismatches 238; Indels 6; Gaps 3;				
QY	1	DAHKSEVAHFKDL---	GREENKALVILAFAYQLQQCPEDFHVKLNVETFAKTVAD	56
Db	28	EGHVDNPHLIGDLTPMIGVDSKGVILAAVSOMPLCPYEHLQVRVEDVMQIDLCAGK	87	
QY	57	ESABENDCKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECTLOHKDDN-PNLFRL	115	
Db	88	ARHANCAKSPMTIILDELCKKPAENAEKVPFHQECCKEDPERHKCFVEHKMANHEELTKY	147	
QY	116	VRPVDVMTAHPDNEETFLKKLYEIAARHPYFYFAPELLFAKRYKAATFTCCQAADKA	175	
Db	148	VRPEQICQDKDAENRGPLLARYITFLAIGHGPHMTIPAILGFAQRDGTIVSHCCKDVEFA	207	
QY	176	A-CLLPKLDLDECKGASSAKORLKASQKQGERAFKAWAVARLSOREPKFAEFAVSKL	234	
Db	208	GQCNDKMPHEKQVEYVCALQKHNCYIIQDPKRALYATYKAVQASOKPPLASFENVOII	267	
QY	235	VTDTKVTTCCHGDLLECADRLAKTICENQDSITSSKKECCCKPLLEKSHCIAEVE	294	
Db	268	VPDVLHQTCCGDMACMLERKMLTAKICKKDELATHLKECCDKPILERSACIILP	327	
QY	295	NDEMPADLSLAADVESKOVKNYAFAKDVPFLGMFIYFYARRHPDYSVVILLRLAKTYE	354	
Db	328	NDOKPADLSKYPHYIDDPVEKLYIEFGGDTWGFGLFYECARRHODYSPMLRLMGSGYE	387	

```

QY 355 TLEKCCARADPHCYAKVDFEPLVEEPONLIKONCELFQOLGEYKFNALLVRYTKK 414
Db 388 EFLKCCCAAGHNECLAETESLKEETESSVTLLKTCGALDKLSYLFQNLIFYVAR 447
QY 415 VPOVSTPTVEYSRNLGKVGKCKHPKAKMPCADYLSVNLQCLVLEHKEKPVSDRYT 474
Db 448 MPALSEGLLRITKSMTWIGKCCRRPDDQMTCSGGIGVFGQICMKQKTFVNEKVA 507
QY 475 KCTESIVNRPCFSALEVDYTPKFNFAETFFHADICTLSEKROIKKQALVELYK 534
Db 508 QCSHLSSTQPCFSALEVDYTPKFNFAETFFHADICTLSEKROIKKQALVELYK 567
QY 535 HKPKAKKEOLKAVMDDEAFVEKCKKADDETCFAEEGKKIYVAASQAALGL 585
Db 568 QYPHTDEQLKTCVNVFVMDQCKADHNECFALGAKLIDACKAILAY 618

RESULT 8
Q8MJ76
ID Q8MJ76 PRELIMINARY; PRT; 610 AA.
AC Q8MJ76
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Alpha-fetoprotein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim J.G., Nonnen D., Vallet J.L., Christenson R.K.;
RL "Mapping of the porcine alpha-fetoprotein (AFP) gene to SSC8.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF517770; AAM66710.1;
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 610 AA; 68624 MW; C985BEAD44963D5E CRC64;

```

```

Query Match 39.3%; Score 1218.5; DB 6; Length 610;
Best Local Similarity 39.7%; Pred. No. 1.9e-87;
Matches 227; Conservative 117; Mismatches 225; Indels 3; Gaps 3;

QY 16 EENFKALVLIAPAYLQCCPFEDHVKLVNVEFPAKTCVADESAENCDKSLHTLFGDK 75
Db 40 EENLVDLATIFFAQFQAEYKQVQKVDLVAVIEKSTGSGPAGCLEQVSVFLEBIC 99
QY 76 TVAILRTYEGMADCCAKQEPERNECFLOHKDDNP-NLPRLVPEVDVMACTAFHDNEFF 134
Db 100 HEERIPKYG-LSEHCSQSGEERHNCFLARKKAAPASIPFQVPEVFTSCAYENRELF 158
QY 135 LKLYLYIARRHPFYFAPPELLFFAKRYKAAPTECCOAAADKACLLPKDLDRDGRKASSA 194
Db 159 MTRVYIARRHPFYFAPPELLFFAKRYKAAPTECCOAAADKACLLPKDLDRDGRKASSA 218
QY 195 KORLKASLQKGERAFKAWAVARSORFPKAEFAVSKLVTDLTQVHTCCGGDLLECA 254
Db 219 LNOHMCVVMRQFARTFRAITVTLSQKFPKANETQIKLVDAVHIEECRCGNVLECL 278
QY 255 DDRADLAKYICENQDSISKLKECCERP-LLEKSHCIAEVENDEMPADLPISLAADFVSK 313
Db 279 QDAERVVYSYCSQDPLSSKTAECCKLPTTLELGSQCIHAENDRPEGLSNLNRFTGER 338
QY 314 DVCNVAEADQFLGMFLYARRHPDYSVVLLLRKAKTYETTLKCCCAADPHCYAKV 373
Db 339 DFNLSREKDLNMAFTYESSRRHPKLVAVPILRVAKGYQELKEKCSQSNPLECDKG 398
QY 374 FDEPKLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKKVPQVSTPTLVEYSRNLGK 433

```

```

Db 399 EEELEKYIOESQALAKRSCGLFQKLGEYLYQNAFLVATKKAQOLTPPELMALTRKMAFT 458
QY 434 GSKCKHPKAKMPCADYLSVNLQCLVLEHKEKPVSDRYTKCTESIVNRPCFSALEV 493
Db 459 GAACCHLSRDLACGGAADLIIGQLCIRHEEMPINPGVGCCTSSYANRRPCFSSIVL 518
QY 494 DETYVPKEPNAETFFHADICTLSEKROIKKQALVELVLEHKEKPAKKEOLKAVMDFAA 553
Db 519 DETYVPKPSDDKFIHFHDKLQCAQVALQPMKQOFLINLVKQKQITEQLEAVIADFSG 578
QY 554 FVEKCKKADDETCFAEEGKKIYVAASQAALGL 585
Db 579 LLEKCCGQGGQEVCFAEEGPALISKTRASLGV 610

RESULT 9
Q8UW06
ID Q8UW06 PRELIMINARY; PRT; 624 AA.
AC Q8UW06
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Serum albumin precursor.
OS Ambystoma texanum (Smallmouth salamander).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidae; Ambystomatidae;
OC Ambystoma.
OX NCBI_TaxID=8304;
RN [1]
RP SEQUENCE FROM N.A.
RA Haverfield E.V., Uzzell T., Spolsky C.M., Bazartseren B.;
RL "Serum albumin of the mole salamanders Ambystoma maculatum and
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217182; AAL56645.1;
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 1.
KW Signal.
FT SIGNAL. 1 24 POTENTIAL.
FT CHAIN 25 624 SERUM ALBUMIN.
SQ SEQUENCE 624 AA; 70321 MW; DE08533BF4953EF7 CRC64;

Query Match 35.0%; Score 1087; DB 13; Length 624;
Best Local Similarity 37.7%; Pred. No. 4.3e-77;
Matches 214; Conservative 102; Mismatches 240; Indels 12; Gaps 5;

QY 14 LGEENFKALVLIAPAYLQCCPFEDHVKLVNVEFPAKTCVADESAENCDKSLHTLFGDK 73
Db 45 LGVEHAKALAMALFSOMLSKCPHHEQVQVRVNMVDIADLCRSAGKHGDCGKSVMTILNE 104
QY 74 LCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNP-NLPRLVPEVDVMACTAFHDNEE 132
Db 105 ICTPENPEKYPFHGECCKEDPERHCKCFIEHSTDPKERTYKVPSPQECQKHAEKND 164
QY 133 TFLAKLYIETARRHPFYFAPPELLFFAKRYKAAPTECCO-AADKAACILPKLDELDRDEKA 191
Db 165 EFLGHYTHKVASSHTMTPPALISFTLHFDGIVSHCCNDEATVGCQLSERMPAKHEVEH 224
QY 192 SSARQRLKASLQKGERAFKAWAVARSORFPKAEFAVSKLVTDLTQVHTCCGGDL 251
Db 225 VCAVQKHNCYILQNFERNALRAKAAHACSKFPFASFENVORLTDGIVHLHQTCGCGDM 284
QY 252 ECADDRADLAKYICENQDSISKLKECCERP-LLEKSHCIAEVENDEMPADLPISLAADFV 311
Db 285 ACMAERKNTTQTCBK-----KKCEKPVLESBCIVLPNDKPAOLSPVRYTFD 336
QY 312 SKDVCKNYAEAKDFLGMFLYARRHPDYSVVLLLRKAKTYETTLKCCCAADPHCYA 371

```

```

Db 337 DPEVCKRFKEGAFMGRLECDYAKIHPEHSAELNRIASGLEKAYKTCAGAHNECIA 396
QY 372 KVEDEKPIVEEPQNLKQNCBELFQOLGEYKFNALLVRYTKVPOVSTPTLVEVSRNLG 431
Db 397 KEETLREHEIASKTLKTKTCGALKGKGYHFQNMIVRYTGLPRSSDAFLITITLT 456
QY 432 KYGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHKEFV-SDRVTKCCTESLVNRRPQFSA 490
Db 457 NIGQCKCKLPEDQPCSGGLGWFAQIC-QNKTTFFENEKLAHCKDLSLFTTPCFAA 515
QY 491 LEVDITYKFEFNAFTTHADICLSEKEROIKKQFALVELVKKHKAQKAVMD 550
Db 516 LTVDETYVPVPAVTAESFNDEFCTPSEADLQAKKQTFMLHLYRTHPTKDQVKTISEK 575
QY 551 FAAFVEKCKADDKETCFABEGKLVAA 578
Db 576 FLAMGQCCQADQNECFATEGAKLVEA 603

```

## RESULT 10

```

Q8BK65 PRELIMINARY; PRT; 605 AA.
ID Q8BK65
AC Q8BK65;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Alpha-fetoprotein precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).
DR EMBL; AK076053; BAC36150.1; -.
SQ SEQUENCE 605 AA; 67322 MW; 049B74A48B01E4B CRC64;

```

```

Query Match 34.9%; Score 1082; DB 11; Length 605;
Best Local Similarity 35.7%; Pred. No. 1e-76;
Matches 204; Conservative 120; Mismatches 241; Indels 6; Gaps 3;

QY 16 EENFKALVLIAPAYLOQCFEDHVKLVNVEYFAKTCVADESAENCDSLTLEGDKLC 75
Db 40 EKNVLSIATITFTQVPEATEEE---VNKMTSDVLAAMKNSGDCLESQLSVFLDEIC 95
QY 76 TVATLRETYGEMADCCAKQEPERNECFLOHKDDNP-NLPRLVPRVDVMTAFHDNETF 134
Db 96 HETELSNKYG-LSGCCSQSGVERHQCLLARKKTAPASVPPFPPEAESCKAHEENRAVF 154
QY 135 LAKYLYETARHPYFYAPPELLFAKRYKAFTCCQADKAACLLPKLDELDECKASSA 194
Db 154 MNRFTYVSRNPNMYPAILLSLAQYDKVVLACCKADNKEECFQTKRASIAREGSM 214
QY 195 KORLACASLOKGEFRAKAWARLSQRPFAEYKSVLTDITKVHTECHGDLLECA 254
Db 155 MNRFTYVSRNPNMYPAILLSLAQYDKVVLACCKADNKEECFQTKRASIAREGSM 214
QY 195 KORLACASLOKGEFRAKAWARLSQRPFAEYKSVLTDITKVHTECHGDLLECA 254
Db 215 LNEHVCSVIRKFSNQLQNTTIKLSQKLTAEANTFQIKLADVAHIEECQGNLECL 274
QY 255 DORADLAKYICENODSISSKLEKCECKPPLKSHCIAEVENDEMPADLPSLAADVSKD 314
Db 215 LNEHVCSVIRKFSNQLQNTTIKLSQKLTAEANTFQIKLADVAHIEECQGNLECL 274
QY 255 DORADLAKYICENODSISSKLEKCECKPPLKSHCIAEVENDEMPADLPSLAADVSKD 314
Db 314 VCKNYAEAKDYVGLMFLYFARHPDYVSVLLRLIAKTYETTLKCCAAADPHECYAKVF 374
QY 315 VCKNYAEAKDYVGLMFLYFARHPDYVSVLLRLIAKTYETTLKCCAAADPHECYAKVF 374
Db 334 FQFSSSEKIMFASFLHEYSRTHPNLPSVILRIAKTYOILEKCSQSGNLPFGQDNLE 394
QY 315 VCKNYAEAKDYVGLMFLYFARHPDYVSVLLRLIAKTYETTLKCCAAADPHECYAKVF 374
Db 335 FQFSSSEKIMFASFLHEYSRTHPNLPSVILRIAKTYOILEKCSQSGNLPFGQDNLE 394
QY 375 DEFKDLVEEPQNLKQNCBELFQOLGEYKFNALLVRYTKVPOVSTPTLVEVSRNLGKV 434

```

```

Db 395 BELQKHIESQALSQSCALYQTLGDYKIQNLQFLIGYTRKAPQLSAELIDLTGKWSIA 454
QY 435 SKCKKHPEAKRMPCAEDYLSVVLNQLCVLHKEFVPSDRVTKCCTESLVNRRPQFSALEVD 494
Db 455 STCCQLEKSEWGGEGMADIFIGHLCIRNEASVNSGISHCNCSYNNRRICTISFLRD 514
QY 495 EYIVPEFNAFTTHADICLSEKEROIKKQFALVELVKKHKAQKAVMDPFAAF 554
Db 515 EYIAPPPSEKFTFHKDLCQAHGKALQTMQELLINLVKQFELTEBQJLAANTADFSGL 574
QY 555 VEKCKADDKETCFABEGKLVAAQAAIGL 585
Db 575 LEKCKAQDQEVCTEFGPKLISKTRDALGV 605

```

## RESULT 11

```

Q8BK56 PRELIMINARY; PRT; 605 AA.
ID Q8BK56
AC Q8BK56;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Alpha-fetoprotein precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Placenta, and Extraembryonic tissue;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).
DR EMBL; AK076197; BAC36249.1; -.
SQ SEQUENCE 605 AA; 67409 MW; CF3509A9EC146110 CRC64;

```

```

Query Match 34.7%; Score 1076; DB 11; Length 605;
Best Local Similarity 35.6%; Pred. No. 3.1e-76;
Matches 203; Conservative 120; Mismatches 242; Indels 6; Gaps 3;

QY 16 EENFKALVLIAPAYLOQCFEDHVKLVNVEYFAKTCVADESAENCDSLTLEGDKLC 75
Db 40 EKNVLSIATITFTQVPEATEEE---VNKMTSDVLAAMKNSGDCLESQLSVFLDEIC 95
QY 76 TVATLRETYGEMADCCAKQEPERNECFLOHKDDNP-NLPRLVPRVDVMTAFHDNETF 134
Db 96 HETELSNKYG-LSGCCSQSGVERHQCLLARKKTAPASVPPFPPEAESCKAHEENRAVF 154
QY 135 LAKYLYETARHPYFYAPPELLFAKRYKAFTCCQADKAACLLPKLDELDECKASSA 194
Db 155 MNRFTYVSRNPNMYPAILLSLAQYDKVVLACCKADNKEECFQTKRASIAREGSM 214
QY 195 KORLACASLOKGEFRAKAWARLSQRPFAEYKSVLTDITKVHTECHGDLLECA 254
Db 215 LNEHVCSVIRKFSNQLQNTTIKLSQKLTAEANTFQIKLADVAHIEECQGNLECL 274
QY 255 DORADLAKYICENODSISSKLEKCECKPPLKSHCIAEVENDEMPADLPSLAADVSKD 314
Db 275 ODGEKVMYICQONILSSKIAECCKLPMIQLGFCIIHAENGKVPGLSNPQFLGDN 334
QY 315 VCKNYAEAKDYVGLMFLYFARHPDYVSVLLRLIAKTYETTLKCCAAADPHECYAKVF 374
Db 335 FQFSSSEKIMFASFLHEYSRTHPNLPSVILRIAKTYOILEKCSQSGNLPFGQDNLE 394
QY 375 DEFKDLVEEPQNLKQNCBELFQOLGEYKFNALLVRYTKVPOVSTPTLVEVSRNLGKV 434
Db 395 BELQKHIESQALSQSCALYQTLGDYKIQNLQFLIGYTRKAPQLSAELIDLTGKWSIA 454
QY 435 SKCKKHPEAKRMPCAEDYLSVVLNQLCVLHKEFVPSDRVTKCCTESLVNRRPQFSALEVD 494

```

```

Db 455 STCCOLSEKWSGCEGMDIFIGHLCIRNEASPVNSGTSHCNSSYSNRRCLITSFLRD 514
QY 495 ETYVPKFEAFNFTPHADICVLSKQKIKKQALVELVKKHKKATKEOLKAYMDDPAAF 554
Db 515 ETYAPPPSDEKFIHKDQOQKQALQTMKQELLINLVKORPELTBEQLAANVADFSGL 574
QY 555 VEKCKADKDETCFABEGKIKLVAAASQALGL 595
Db 575 LEKCKAQDQEVCTEFGPKLISKTRDALGV 605

RESULT 12
Q8JIA7
ID Q8JIA7 PRELIMINARY; PRT; 400 AA.
AC Q8JIA7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE A/B over-sized serum albumin (fragment).
OS Sphenodon punctatus (Hatteria) (Tuatara).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosaunia; Sphenodontia; Sphenodontidae; Sphenodon.
OX NCBI_TaxID=8508;
RN [1]
RP SEQUENCE FROM N.A.
RA Metcalf V.J., Brennan S.O., George P.M., Chambers G.K.;
RT "Partial mRNA sequence for tuatara A/B serum albumin."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF375973; AAM46106.1; -
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM0103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
FT NON_TER 1 1
FT SIGNAL <1 23
FT CHAIN 24 603 SERUM ALBUMIN.
FT SIGNAL 24 603
SQ SEQUENCE 400 AA; 45715 MW; 8DE206096570CF753 CRC64;

Query Match 33.7%; Score 1045; DB 13; Length 400;
Best Local Similarity 47.7%; Pred. No. 5.1e-74;
Matches 186; Conservative 80; Mismatches 124; Indels 0; Gaps 0;

QY 196 QELKASIQKFGRAKAWARLARSORPFAEAVSKLVTLTKVHTTECHGDLLECAD 255
Db 3 EKHSGLFKSGRAFOAKLVQLSQKFPKAPFEELHKVVLVATKLQKCECHGDMIECLD 62
QY 256 DRADLAKYICENQDSISSKLECKEPLLEKSHCIAEVENDEMPADLPSLAADFVESKDV 315
Db 63 DRVEVMAYICSKQAVFSSKIDCKCEKPIVDREVCILQADLDEKPADLPSIAGQYIESTEV 122
QY 316 CENYAEAKDVLGNFLYEYARRHPDYSVVLLRLAKTYETTELKCCAAADPHECYAKVD 375
Db 123 KKHVEGKDVLAHFVYEYSRHPFESSQMLLRGTQYDLDKCKTENPEPCYKAGE 182
QY 376 EFKPLVEEPONLKKONCELFGEYKFNALVRRTKKVPQVSTPTLVEVSRLNGKVG 435
Db 183 ELARIHQESQELKTHCSFTYTSQGDQPFQKMLVRYTKMPQLPAEELIETSKLTGVG 242
QY 436 KCKHPKAKMPCADYLSVLNQLCVLHEKTPVSDRVYKCTCESLVRNRCFCFSALVEVD 495
Db 243 KCCPLSEDKKLSSEKLSWLFECRCQHEASPVNNHVHCCCTDSYSEMRFCFTKLGVDD 302
QY 496 TVVPKEFNAFTFHADICTLSEKROIKKOTALVELVKKHKKATKEOLKAYMDDPAAFV 555
Db 303 SVTPEFCPSFTFLDEQLCTAPEARLKKQLTFLVLKLIQLKPOIEQLKLVTDYHAME 362
QY 556 EKCKKADKDETCFAEKGKILVAASQALGL 585
Db 363 EKCCQAKNQCFCFTEGKLTQEGKALLGV 392

RESULT 13

```

```

Q9YGH6
ID Q9YGH6 PRELIMINARY; PRT; 603 AA.
AC Q9YGH6;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Serum albumin precursor (fragment).
GN ALB.
OS Rana shqiperica.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
OX NCBI_TaxID=44326;
RN [1]
RP SEQUENCE FROM N.A.
RA Uzzell T., Hotz H.;
RT "Albumin cDNA sequence of Rana shqiperica: evolutionary changes in frog albumins."
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40452; AAD09358.1; -
DR HSSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM0103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Signal.
FT NON_TER 1 1
FT SIGNAL <1 23
FT CHAIN 24 603 SERUM ALBUMIN.
FT SIGNAL 24 603
SQ SEQUENCE 603 AA; 69293 MW; 340D3733FA010C99 CRC64;

Query Match 30.8%; Score 955; DB 13; Length 603;
Best Local Similarity 33.6%; Pred. No. 9.9e-67;
Matches 190; Conservative 119; Mismatches 241; Indels 16; Gaps 3;

QY 12 KDIGENFKALVLIAPAOYLQOCPFDHVKLVNVEYTFEAKTCVADSAENCKSLHTLFG 71
Db 37 KAVGKFAVEKLVLMVAQDFEKLSDHLKVOAKIEADVNCNEKHPAEKCKPAEILYH 96
QY 72 DKLCVTATREYVEMADCCAKQPERNECFQKHDDNPNLPLVR-----PEVDVMT 125
Db 97 DIVCKEEDIDQLYPTWTECCGKAERTKCFYEHE-----VRVEEKIPNIESCK 148
QY 126 AFHNEETTKYLYEYIARRHPYAPPELLFPKRYKAAPTECCOAAADKAACLPKLDEL 185
Db 149 EHKEHFQRAFSTYLSNIAKRHSKLYPPAIVLGAIQYNEITTECCAAEDKAKCFGERMPQV 208
QY 186 RDEKASSAKQRLKASQKFGERAFAKAVARLARSORPFAEAVSKLVTLTKVHTEC 245
Db 209 KLLTNVLEDKHKQCEVLEKPEPVSQALTVQVSORFGNAKYDDVEKVTIEAHLNEDC 268
QY 246 CHGDLFECADRADLAKYICENQDSISSKLECKEPLLEKSHCIAEVENDEMPADLPSL 305
Db 269 CKGDVETCMEATEATEICIAKEKLSKLSDDCAKGVLETPCILALPNEE--PDLPTE 326
QY 306 AADPVESKDVCKNYAEAKDVLGNFLYEYARRHPDYSVVLLRLAKTYETTELKCCAAAD 365
Db 327 LKEYIEDVHCYENQDRARKTYLAHTFIDYSRHSQSSQSCLVNRSRGFEMLEKCCASAN 386
QY 366 PHECYAKVDFECPKLVPEPONLKKONCELFGEYKFNALVRRTKKVPQVSTPTLVE 425
Db 387 SAECLDKAPLLEAALKENEETISKQNGALEKLGFNDFVIQLLVRYRKMPQVTAQTIVE 446
QY 426 VSRNLGKVSCKCKPIEAKRMPCADYLSVLNQLCVLHEKTPVSDRVYKCTCESLVRN 485
Db 447 LTGMAKIGYVCCGLPDNKKQPCAEKLDLIGEMCEKEKKTFFINDRVHCCYDSTANRR 506
QY 486 PCFSALFVDRTPVYKPEFNAFTFHADICTLSEKROIKKOTALVELVKKHKKATKEOLK 545
Db 507 PCFTKGPYANPEAPVWDESKLHFTADMCKGSADQLKTLVLVLEFKMRKPTCGREKLT 566

```

[illegible]

QY 532 LVKHKPKATKEQLKAVMDPAAFVEKCKKADDKETCFAEKGKLVAAASQAL 583  
:||| || | |||||: || ||||| :| || || ||| :|||  
Db 354 VVKCKPAITHEQLKAVITDFYGVVEKCKCHGENHEACFLARGPQLVQFTQAL 405

Search completed: August 31, 2003, 16:36:34  
Job time : 111 secs